



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년06월26일
(11) 등록번호 10-2547738
(24) 등록일자 2023년06월21일

(51) 국제특허분류(Int. Cl.)
C07K 16/00 (2006.01) C07K 16/46 (2006.01)
C12N 15/90 (2006.01) C12N 5/071 (2010.01)
(52) CPC특허분류
C07K 16/00 (2013.01)
C07K 16/468 (2013.01)
(21) 출원번호 10-2018-7029365
(22) 출원일자(국제) 2017년04월20일
심사청구일자 2020년04월14일
(85) 번역문제출일자 2018년10월11일
(65) 공개번호 10-2018-0134893
(43) 공개일자 2018년12월19일
(86) 국제출원번호 PCT/US2017/028555
(87) 국제공개번호 WO 2017/184832
국제공개일자 2017년10월26일
(30) 우선권주장
62/325,400 2016년04월20일 미국(US)
(56) 선행기술조사문헌
WO2008151219 A1*
KR1020150110571 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
리제너론 파마슈티칼스 인코포레이티드
미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리버 로드 777
(72) 발명자
밥 로버트
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
부라코프 다리야
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내
(뒷면에 계속)
(74) 대리인
특허법인와이에스장

전체 청구항 수 : 총 40 항

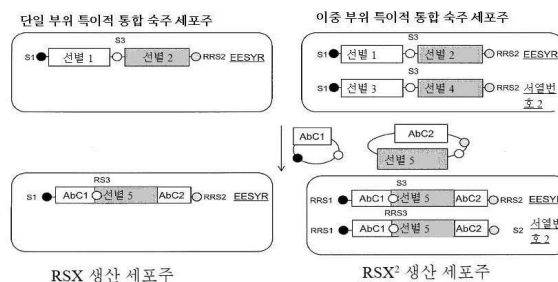
심사관 : 문영준

(54) 발명의 명칭 발현 강화 유전자좌의 사용에 기초하여 항체를 만들기 위한 조성물 및 방법

(57) 요약

본 발명은 진핵 세포에서 재조합 단백질을 부위 특이적으로 통합하고 및 발현시키는 것에 관한 것이다. 특히, 본 발명은 다중 발현 강화 유전자좌를 사용하여 진핵 세포, 특히 차이니스 햄스터(*Cricetulus griseus*) 세포주에서 (단일 특이적 항체 및 이중 특이적 항체를 포함하는) 항원 결합 단백질의 발현을 개선하기 위한 조성물 및 방법에 관한 것이다.

대표도



(52) CPC특허분류

C12N 15/907 (2013.01)

C12N 5/0602 (2013.01)

C07K 2317/31 (2013.01)

C07K 2317/526 (2013.01)

C12N 2510/02 (2013.01)

(72) 발명자

첸 강

미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버
로드 777 리제너론 파마슈티칼스 인코포레이티드
내

팬들 제임스 피.

미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버
로드 777 리제너론 파마슈티칼스 인코포레이티드
내

자오 유

미국 뉴욕 11596 윌리스턴 파크 홀리데이 파크 드
라이브 14

명세서

청구범위

청구항 1

제1 발현 강화 유전자와 내에 통합된 제1 외인성 핵산; 및 제2 발현 강화 유전자와 내에 통합된 제2 외인성 핵산을 포함하는 세포로서,

상기 2개의 발현 강화 유전자와 중 하나는 서열번호 1의 뉴클레오티드 서열을 포함하고, 상기 2개의 발현 강화 유전자와 중 다른 하나는 서열번호 2 또는 서열번호 3의 뉴클레오티드 서열을 포함하고,

상기 제1 및 제2 외인성 핵산은 함께 항원 결합 단백질을 암호화하는, 세포.

청구항 2

제1항에 있어서, 상기 제1 외인성 핵산은 제1 중쇄 단편(HCF)을 암호화하는 뉴클레오티드 서열을 포함하고, 상기 제2 외인성 핵산은 제1 경쇄 단편(LCF)을 암호화하는 뉴클레오티드 서열을 포함하는, 세포.

청구항 3

제2항에 있어서, 상기 항원 결합 단백질은 단일 특이적인, 세포.

청구항 4

제2항에 있어서, 상기 제2 외인성 핵산은 제2 HCF를 암호화하는 뉴클레오티드 서열을 추가로 포함하는, 세포.

청구항 5

제4항에 있어서,

a) 상기 제1 및 제2 HCF는 상이하며; 또는

b) 상기 제1 HCF를 암호화하는 상기 뉴클레오티드 서열은 제1 CH3 도메인을 암호화하고, 상기 제2 HCF를 암호화하는 상기 뉴클레오티드 서열은 제2 CH3 도메인을 암호화하는,

세포.

청구항 6

제5항에 있어서, a)에서: 상기 항원 결합 단백질은 이중 특이적인, 세포.

청구항 7

제5항에 있어서, b)에서:

i) 상기 제1 및 제2 CH3 도메인은 적어도 하나의 아미노산 위치가 상이하거나; 또는

ii) 상기 제1 및 제2 CH3 도메인을 암호화하는 상기 뉴클레오티드 서열은 상기 뉴클레오티드 서열 중 하나가 코돈 변형되었다는 점이 서로 상이한,

세포.

청구항 8

제4항에 있어서, 상기 제1 외인성 핵산은 제2 LCF를 암호화하는 뉴클레오티드 서열을 추가로 포함하는, 세포.

청구항 9

제8항에 있어서, 상기 제1 및 제2 LCF는 동일한, 세포.

청구항 10

제2항에 있어서, 제1 재조합 효소 인식 부위 (RRS) 및 제2 RRS는 상기 제1 발현 강화 유전자좌 내에 통합된 제1 외인성 핵산에 대해 상대적으로 5' 및 3'에 각각 위치되고, 제3 RRS 및 제4 RRS는 상기 제2 발현 강화 유전자좌 내에 통합된 제2 외인성 핵산에 대해 상대적으로 5' 및 3'에 각각 위치되고, 상기 제1 및 제2 RRS가 상이하고, 상기 제3 및 제4 RRS가 상이한, 세포.

청구항 11

제10항에 있어서, 상기 제1, 제2, 제3 및 제4 RRS는 서로 상이한, 세포.

청구항 12

제4항에 있어서, 제1 추가 재조합 효소 인식 부위 (RRS)가 상기 제1 LCF를 암호화하는 상기 뉴클레오타이드 서열과 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열 사이에 존재하는, 세포.

청구항 13

제8항에 있어서, 제1 재조합 효소 인식 부위 (RRS) 및 제2 RRS는 상기 제1 발현 강화 유전자좌 내에 통합된 제1 외인성 핵산에 대해 상대적으로 5' 및 3'에 각각 위치되고, 제3 RRS 및 제4 RRS는 상기 제2 발현 강화 유전자좌 내에 통합된 제2 외인성 핵산에 대해 상대적으로 5' 및 3'에 각각 위치되고, 상기 제1 및 제2 RRS가 상이하고, 상기 제3 및 제4 RRS가 상이한, 세포.

청구항 14

제13항에 있어서,

- a) 상기 제1 및 제2 HCF가 동일하고, 상기 제1 및 제2 LCF가 동일하고, 상기 제1 및 제3 RRS가 동일하고, 상기 제2 및 제4 RRS가 동일하고; 또는
- b) 상기 제1 및 제2 HCF가 상이하고, 상기 제1 및 제2 LCF가 동일하고, 상기 제1, 제2, 제3 및 제4 RRS가 서로 상이한, 세포.

청구항 15

제13항에 있어서, 제1 추가 RRS가 상기 제2 외인성 핵산 내에서 상기 제1 LCF를 암호화하는 상기 뉴클레오타이드 서열과 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열 사이에 존재하고, 상기 제1 추가 RRS는 상기 제1, 제2, 제3 및 제4 RRS와 상이한, 세포.

청구항 16

제15항에 있어서, 제2 추가 RRS가 상기 제2 LCF를 암호화하는 상기 뉴클레오타이드 서열과 상기 제1 HCF를 암호화하는 상기 뉴클레오타이드 서열 사이에 존재하고, 상기 제1 및 제2 추가 RRS는 동일하거나 상이하고 각각 상기 제1, 제2, 제3 및 제4 RRS와 상이한, 세포.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, 상기 세포는 CHO 세포인, 세포.

청구항 18

세포 및 벡터 세트를 포함하는 시스템으로서,

상기 세포는

제1 발현 강화 유전자좌 내에 통합된 제1 재조합 효소 인식 부위 (RRS), 제1 외인성 핵산 및 제2 RRS(5'에서 3'까지); 및

제2 발현 강화 유전자좌 내에 통합된 제3 RRS, 제2 외인성 핵산 및 제4 RRS(5'에서 3'까지)를 포함하고,

상기 제1 및 제2 RRS가 상이하고 상기 제3 및 제4 RRS가 상이하고, 상기 2개의 발현 강화 유전자좌 중 하나는 서열번호 1의 뉴클레오타이드 서열을 포함하고, 상기 2개의 발현 강화 유전자좌 중 다른 하나는 서열번호 2 또는 서열번호 3의 뉴클레오타이드 서열을 포함하고,

상기 벡터 세트는

제1 벡터 5' RRS, 제1 핵산, 및 제1 벡터 3' RRS(5'에서 3'까지)를 포함하고 상기 제1 벡터 5' 및 3' RRS는 상이한 제1 벡터;

제2 벡터 5' RRS, 제2 핵산, 및 제2 벡터 3' RRS(5'에서 3'까지)를 포함하고 상기 제2 벡터 5' 및 3' RRS는 상이한 제2 벡터; 및

제1 중쇄 단편 (HCF)를 암호화하는 뉴클레오티드 서열 및 제1 경쇄 단편 (LCF)를 암호화하는 뉴클레오티드 서열을 포함하며, 상기 뉴클레오티드 서열 중 하나는 상기 제1 핵산 내에 있고 다른 하나의 뉴클레오티드 서열은 상기 제2 핵산 내에 있고, 상기 제1 HCF와 상기 제1 LCF는 항원 결합 단백질의 영역이고,

상기 세포 내로 상기 벡터가 도입될 때, 상기 벡터 내의 상기 제1 및 제2 핵산은 상기 RRS에 의해 매개된 재조합을 통해 상기 제1 발현 강화 유전자와 및 상기 제2 발현 강화 유전자와 내로 각각 통합되는, 시스템.

청구항 19

제18항에 있어서, 상기 항원 결합 단백질은 단일 특이적 항원 결합 단백질인, 시스템.

청구항 20

제19항에 있어서,

a) 상기 제1 및 제3 RRS가 동일하고, 상기 제2 및 제4 RRS가 동일하며; 또는

b) 상기 제1 벡터 5' RRS가 상기 제1 RRS와 동일하고, 상기 제1 벡터 3' RRS가 상기 제2 RRS와 동일하고; 상기 제2 벡터 5' RRS가 상기 제3 RRS와 동일하고, 상기 제2 벡터 3' RRS가 상기 제4 RRS와 동일한, 시스템.

청구항 21

제20항에 있어서, a)에서: 제1 추가 RRS가 상기 제1 유전자좌에서 상기 제1 및 제2 RRS 사이에 존재하는, 시스템.

청구항 22

제21항에 있어서, 상기 제1 벡터 5' RRS가 상기 제1 및 제3 RRS와 동일하고; 상기 제1 벡터 3' RRS, 상기 제2 벡터 5' RRS, 및 상기 제1 추가 RRS가 동일하고; 상기 제2 벡터 3' RRS가 상기 제2 및 제4 RRS와 동일한, 시스템.

청구항 23

제22항에 있어서, 상기 LCF 암호화 뉴클레오티드 서열은 상기 제1 벡터 내에 있고, 상기 HCF 암호화 뉴클레오티드 서열은 상기 제2 벡터 내에 있는, 시스템.

청구항 24

제18항에 있어서, 상기 항원 결합 단백질은 이중 특이적 항원 결합 단백질인, 시스템.

청구항 25

제24항에 있어서, 상기 제1 HCF와 상이한 제2 HCF를 암호화하는 뉴클레오티드 서열을 추가로 포함하는, 시스템.

청구항 26

제25항에 있어서,

i) 상기 제1 LCF를 암호화하는 상기 뉴클레오티드 서열과 상기 제2 HCF를 암호화하는 상기 뉴클레오티드 서열 모두가 상기 제1 벡터 내의 상기 제1 핵산에 포함되고, 상기 제1 HCF를 암호화하는 상기 뉴클레오티드 서열은 상기 제2 벡터 내에 있으며; 또는

ii) 상기 제2 HCF를 암호화하는 상기 뉴클레오티드 서열은 제3 벡터 5' RRS 및 제3 벡터 3' RRS가 측면에 위치한 별도의 제3 벡터 상에 있으며; 또는

iii) 상기 제1 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제1 CH3 도메인을 암호화하고, 상기 제2 HCF를 암호화하는 상기 뉴클레오타이드 서열은 제2 CH3 도메인을 암호화하는,

시스템.

청구항 27

제26항에 있어서, i)에서: 상기 제1 벡터 5' RRS가 상기 제1 RRS와 동일하고, 상기 제1 벡터 3' RRS가 상기 제2 RRS와 동일하고, 상기 제2 벡터 5' RRS가 상기 제3 RRS와 동일하고, 상기 제2 벡터 3' RRS가 상기 제4 RRS와 동일한, 시스템.

청구항 28

제26항에 있어서, ii)에서: 상기 제1 LCF를 암호화하는 상기 뉴클레오타이드 서열은 상기 제1 벡터 내에 있고, 상기 제1 HCF를 암호화하는 상기 뉴클레오타이드 서열은 상기 제2 벡터 내에 있고, 상기 제1 벡터 5' RRS가 상기 제1 RRS와 동일하고, 상기 제1 벡터 3' RRS가 상기 제2 벡터 5' RRS 및 제1 추가 RRS와 동일하고, 상기 제2 벡터 3' RRS가 상기 제2 RRS와 동일하고, 상기 제3 벡터 5' RRS가 상기 제3 RRS와 동일하고, 상기 제3 벡터 3' RRS가 상기 제4 RRS와 동일하고, 상기 제1 추가 RRS는 상기 제1 및 제2 RRS 사이의 상기 제1 유전자좌에 포함되는, 시스템.

청구항 29

제26항에 있어서, iii)에서:

- a) 상기 제1 및 제2 CH3 도메인은 적어도 하나의 아미노산이 상이하거나; 또는
- b) 상기 제1 및 제2 CH3 도메인을 암호화하는 상기 뉴클레오타이드 서열은 상기 뉴클레오타이드 서열 중 하나가 코돈 변형되었다는 점이 상이한,

시스템.

청구항 30

제27항에 있어서, 상기 벡터의 세트는 제2 LCF를 암호화하는 뉴클레오타이드 서열을 추가로 포함하는, 시스템.

청구항 31

제30항에 있어서, 상기 제1 및 제2 LCF는 동일한, 시스템.

청구항 32

제30항에 있어서, 상기 제2 LCF를 암호화하는 상기 뉴클레오타이드 서열은

- a) 상기 제2 벡터의 상기 제2 핵산 내에 있고, 상기 제1 벡터 5' RRS는 상기 제1 RRS와 동일하고, 상기 제1 벡터 3' RRS는 상기 제2 RRS와 동일하고, 상기 제2 벡터 5' RRS는 상기 제3 RRS와 동일하고, 상기 제2 벡터 3' RRS는 상기 제4 RRS와 동일하고; 또는

- b) 제3 벡터 5' RRS 및 제3 벡터 3' RRS가 측면에 위치한 별도의 제3 벡터 상에 있는,

시스템.

청구항 33

제32항에 있어서, b)에서: 상기 제1 벡터 5' RRS 및 3' RRS는 상기 제1 유전자좌에서의 상기 제1 RRS 및 제2 RRS와 각각 동일하고; 상기 제3 벡터 5' RRS는 상기 제3 RRS와 동일하고, 상기 제3 벡터 3' RRS는 상기 제2 벡터 5' RRS 및 상기 제2 유전자좌에서의 상기 제3 및 제4 RRS 사이에 존재하는 추가 RRS와 동일하고, 상기 제2 벡터 3' RRS는 상기 제4 RRS와 동일한, 시스템.

청구항 34

제18항 내지 제33항 중 어느 한 항에 있어서, 상기 제1 HCF를 암호화하는 뉴클레오타이드 서열 및 제1 LCF를 암호

화하는 뉴클레오티드 서열은 서로 프로모터에 작동 가능하게 연결되는, 시스템.

청구항 35

제18항 내지 제33항 중 어느 한 항에 있어서, 상기 세포는 CHO 세포인, 시스템.

청구항 36

(i) 제18항 내지 제33항 중 어느 한 항의 시스템을 제공하는 단계;

(ii) 상기 벡터를 상기 세포 내로 형질감염에 의해 도입하는 단계; 및

(iii) 상기 벡터 내의 상기 핵산이 상기 RRS에 의해 매개된 재조합을 통해 상기 제1 및 제2 발현 강화 유전자좌 내에 통합된 형질감염 세포를 선별하는 단계를 포함하는, 방법.

청구항 37

제36항에 있어서,

(iv) 상기 선별된 형질감염 세포로부터 상기 항원 결합 단백질을 발현시키고 수득하는 단계를 추가로 포함하는, 방법.

청구항 38

제36항에 있어서, 상기 세포는 CHO 세포인, 방법.

청구항 39

항원 결합 단백질을 만드는 방법으로서, 제1항 내지 제16항 중 어느 한 항의 세포를 제공하는 단계, 및 상기 세포로부터 상기 항원 결합 단백질을 발현시키고 수득하는 단계를 포함하는, 방법.

청구항 40

제39항에 있어서, 상기 세포는 CHO 세포인, 방법.

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호 참조

[0002] 본 출원은 그 내용 전체가 참조로서 본원에 통합되는, 2016년 4월 20일에 출원된 미국 가출원 번호 제 62/325,400호의 우선권의 이익을 주장한다.

[0003] 기술분야

[0004] 본 개시는 진행 세포에서의 재조합 단백질의 부위 특이적 통합 및 발현에 관한 것이다. 특히, 본 개시는 발현 강화 유전자좌를 사용함으로써 진행 세포, 특히 차이니스 햄스터(*Cricetulus griseus*) 세포주에서 (단일 특이적 항체 및 이중 특이적 항체를 포함하는) 항원 결합 단백질의 발현을 개선하는 위한 조성물 및 방법에 관한 것이다.

배경 기술

[0005] 세포 발현 시스템의 목적은, 연구용이거나 치료용이거나를 불문하고, 특정 단백질의 제조를 위한 신뢰할 수 있고 효율적인 공급원을 제공하는 것이다. 포유류 세포에서의 재조합 단백질을 발현시키는 것은 치료용 단백질 제조를 위한 바람직한 방법인데, 이는, 예를 들어, 재조합 단백질을 적절하게 번역한 후 변형하는 포유류 발현 시스템의 능력에 기인한다.

[0006] 다양한 발현 시스템을 이용할 수 있음에도 불구하고, 재조합 단백질의 발현을 위한 효율적인 유전자 전달 및 통합 유전자의 안정성에 대한 어려움은 여전히 존재한다. 표적 전이 유전자의 장기적인 발현을 위해서는, 세포주의 표현형이 변하지 않도록 세포 유전자의 파괴를 최소화하는 것을 고려할 수 있다.

[0007] 안정적인 세포주를 조작하여 발현을 위한 다중 유전자, 예컨대 다중 특이적 항체에서와 같이 다중 항체 사슬을 수용하는 것은 특히 어렵다. 통합된 유전자의 발현 수준은 매우 다양할 수 있다. 유전자를 추가로 통합하면 국부적인 유전적 환경(즉, 자리 효과(position effect))으로 인해 발현의 다양성과 불안정성이 더 커질 수 있다. 다중 특이적 항원 결합 단백질의 생산을 위한 발현 시스템은 종종 특정 다중 결합 포맷으로 쌍을 이루도록 의도된 둘 이상의 상이한 면역 글로불린 사슬의 발현을 필요로 하고, 종종 원하는 이중 이량체(heterodimer) 또는 다량체(multimer) 조합보다는 동종 이량체의 생산에 유리하게 작용할 수 있다. 따라서, 당업계에는 개선된 포유류 발현 시스템에 대한 요구가 있다.

발명의 내용

[0008] 일 양태에서, 본 개시는 다수의 외인성 핵산이 함께 항원 결합 단백질을 암호화하는 2개의 발현 강화 유전자좌에서 부위 특이적으로 통합된 다수의 외인성 핵산을 함유하는 세포를 제공한다. 항원 결합 단백질은 이중 특이적 항원 결합 단백질 또는 통상의 단일 특이적 항원 결합 단백질일 수 있다.

[0009] 일부 구현예에서, 제1 강화 발현 유전자좌 내에 통합된 제1 외인성 핵산 및 제2 강화 발현 유전자좌 내에 통합된 제2 외인성 핵산을 함유하는 세포가 제공되며, 제1 및 제2 외인성 핵산은 함께 항원 결합 단백질을 암호화한다.

[0010] 일부 구현예에서, 제1 외인성 핵산은 제1 중쇄 단편(HCF)을 암호화하는 뉴클레오타이드 서열을 함유하고, 제2 외인성 핵산은 제1 경쇄 단편(LCF)을 암호화하는 뉴클레오타이드 서열을 함유한다.

[0011] 일부 구현예에서, 제2 외인성 핵산은 제2 HCF(및 HCF*)를 암호화하는 뉴클레오타이드 서열을 추가로 포함한다. 제1 및 제2 HCF는 이중 특이적 항원 결합 단백질에 있어서와 마찬가지로 동일하거나 상이할 수 있다. 각각의 HCF 암호화 뉴클레오타이드 서열 또는 LCF 암호화 뉴클레오타이드 서열은 불변 영역으로부터 아미노산을 암호화할 수 있다. 일부 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 CH3 도메인을 암호화하고, 제2 HCF(HCF*)를 암호화하는 뉴클레오타이드 서열은 제2 CH3 도메인을 암호화한다. 일부 구현예에서, 제1 및 제2 CH3 도메인은 적어도 하나의 아미노산 위치, 예컨대 상이한 단백질 A 결합 특성을 유도하는 위치가 상이할 수 있다. 다른 구현예에서, 제1 및 제2 CH 도메인을 암호화하는 뉴클레오타이드 서열은 뉴클레오타이드 서열 중 하나가 코돈 변형되었다는 점이 서로 상이하다.

- [0012] 일부 구현예에서, (제1 HCF 암호화 뉴클레오타이드 서열을 함유하는) 제1 외인성 핵산은 제2 LCF를 암호화하는 뉴클레오타이드 서열을 추가로 포함한다. 제2 LCF는 제2 외인성 핵산 서열 내의 제1 LCF와 동일하거나 상이할 수 있다.
- [0013] 본원에서 제공된 세포에 대한 많은 구현예에서, HCF 또는 LCF를 암호화하는 뉴클레오타이드 서열 각각은 독립적으로 프로모터에 작동 가능하게 연결됨으로써, 각각의 HCF 암호화 서열 또는 LCF 암호화 서열의 전사가 개별적으로 조절된다.
- [0014] 일부 구현예에서, 제1 RRS 및 제2 RRS는 제1 외인성 핵산에 대해 상대적으로 5' 및 3'에 각각 위치되고, 제3 RRS 및 제4 RRS는 제2 외인성 핵산에 대해 상대적으로 5' 및 3'에 각각 위치되되, 제1 RRS와 제2 RRS가 상이하고, 제3 RRS와 제4 RRS가 상이하다. 일반적으로, 외인성 핵산의 측면에 위치하는 한 쌍의 RRS 내의 RRS는 핵산의 의도되지 않은 재조합 및 제거를 피하기 위해 상이하다. 일부 구현예에서, 제1, 제2, 제3 및 제4 RRS는 모두 서로 상이하다.
- [0015] 제1 유전자와 내의 제1 외인성 핵산이 제1 HCF 암호화 뉴클레오타이드 서열을 포함하고, 제2 유전자와 내의 제2 외인성 핵산이 제2 LCF 암호화 서열 및 제2 HCF 암호화 서열을 포함하는 구현예에서, 제1 추가 RRS가 제1 LCF를 암호화하는 뉴클레오타이드 서열 및 제2 HCF를 암호화하는 뉴클레오타이드 서열 사이에 존재할 수 있다. 추가 RRS는 제1, 제2, 제3 및 제4 RRS 각각과 상이할 수 있다. 일부 구현예에서, 제1 추가 RRS가 선별 표지 유전자가 작동 가능하게 연결되는 프로모터와 선별 표지 유전자 사이에 포함되거나, 제1 LCF 암호화 서열과 HCF 암호화 서열 사이 또는 제1 HCF 암호화 서열과 제2 HCF 암호화 서열 사이에 존재하는 추가 RRS가 선별 표지 유전자에 또는 선별 표지 유전자의 인트론 내에 포함될 수 있다.
- [0016] 제1 유전자와 내의 제1 외인성 핵산이 제1 HCF 암호화 뉴클레오타이드 서열 및 제2 HCF 암호화 뉴클레오타이드 서열을 포함하고, 제2 유전자와 내의 제2 외인성 핵산이 제1 LCF 암호화 서열 및 제2 HCF 암호화 서열 모두를 포함하는 구현예에서, 제1 RRS 및 제2 RRS는 제1 외인성 핵산에 대해 상대적으로 5' 및 3'에서 각각 존재할 수 있고, 제3 RRS 및 제4 RRS는 제2 외인성 핵산에 대해 상대적으로 5' 및 3'에서 각각 존재할 수 있으며, 제1 RRS와 제2 RRS가 상이하고, 제3 RRS와 제4 RRS가 상이하다. 일부 구현예에서, 제1 RRS와 제3 RRS가 동일하고, 제2 RRS와 제4 RRS가 동일하도록 RRS가 조작될 수 있는 경우, 제1 HCF와 제2 HCF는 동일하고, 제1 LCF와 제2 LCF도 동일하다. 일부 구현예에서, 제1, 제2, 제3 및 제4 RRS가 모두 서로 상이하도록 RRS가 조작될 수 있는 경우, 제1 HCF와 제2 HCF는 상이하고, 제1 LCF와 제2 LCF는 동일하다. 두 HCF의 동일 또는 상이 여부와 무관하게, 추가 RRS는 제1 LCF 암호화 서열과 제2 HCF 암호화 서열 사이에 존재할 수 있고/있거나 제2 LCF 암호화 서열과 제1 HCF 암호화 서열 사이에 존재할 수 있다. 추가 (중간) RRS는 제1, 제2, 제3 및 제4 RRS 각각과 상이하다. 추가 RRS는 두 HCF/LCF 코딩 서열 사이에 위치한 선별 표지 유전자 내에 포함되거나, 선별 표지 유전자의 인트론 내에 포함될 수 있다.
- [0017] 또 다른 양태에서, 두 개의 발현 강화 유전자와 내에 통합된 여러 쌍의 RRS를 함유하고, RMCE를 통해, 항원 결합 단백질을 암호화하는 핵산을 통합하는 데 사용할 수 있는 세포가 제공된다.
- [0018] 특정 구현예에서, 두 개의 발현 강화 유전자와 내에 통합된 여러 쌍의 RRS를 함유하고, 재조합 효소의 존재 하에 RMCE를 통해, 항원 결합 단백질을 암호화하는 핵산을 동시에 통합하는 데 사용될 수 있는 세포가 제공된다.
- [0019] 일부 구현예에서, 5'에서 3'까지 제1 강화된 발현 유전자와 내에 통합된 제1 RRS, 제1 외인성 핵산 및 제2 RRS, 및 5'에서 3'까지 제2 강화 발현 유전자와 내에 통합된 제3 RRS, 제2 외인성 핵산, 및 제4 RRS를 포함하되, 제1 RRS와 제2 RRS는 상이하고 제3 RRS와 제4 RRS는 상이한, 세포가 제공된다.
- [0020] 일부 구현예에서, 제1 외인성 핵산은 제1 선별 표지 유전자를 포함하고, 제2 외인성 핵산은 제2 선별 표지 유전자를 포함하며, 제1 및 제2 선별 표지 유전자는 상이하다.
- [0021] 일부 구현예에서, 제1 및 제2 외인성 핵산 중 하나 또는 둘 다는 추가 RRS를 포함할 수 있는데, 구체적으로, 제1 유전자와 내의 제1 RRS와 제2 RRS 사이에 추가 RRS를 포함하고/하거나 제3 RRS와 제4 RRS 사이에 추가 RRS를 포함할 수 있다. 추가적인 중간 RRS는 5' 및 3'에서의 RRS와 상이하다. 추가 RRS가 5' RRS와 3' RRS 사이에 (예를 들어, 제1 RRS와 제2 RRS 사이에) 포함되는 경우, 하나의 선별 표지 유전자가 5' RRS와 추가 (중간) RRS 사이에 포함될 수 있고, 또 다른 상이한 선별 표지 유전자가 추가 RRS와 3' RRS 사이에 포함될 수 있다.
- [0022] 또 다른 구현예에서, 세포는 제3 RRS(즉, 추가적인 RRS)를 제1 유전자와 내의 제1 RRS와 제2 RRS 사이에 포함하는 제1 외인성 핵산을 제공하는데, 제1 및 제2 RRS는 발현 카세트의 5' 및 3' 말단에서 두 개의 선별 표지의 측

면에 위치한다. 다른 구현예에서, 제2 외인성 핵산도 동일한 제3 RRS(즉, 추가적인 RRS)를 제2 유전자와 내의 제1 RRS와 제2 RRS 사이에 포함할 수 있는데, 제1 및 제2 RRS는 발현 카세트의 5' 및 3' 말단에서 두 개의 선별 표지의 측면에 위치한다. 제1 및 제2 RRS 사이에 포함된 4개의 선별 표지 유전자는 서로 상이하다.

[0023] 또 다른 구현예에서, 세포는 제3 RRS(즉, 추가적인 RRS)를 제1 유전자와 내의 제1 RRS와 제2 RRS 사이에 포함하는 제1 외인성 핵산을 제공하는데, 제1 및 제2 RRS는 발현 카세트의 5' 및 3' 말단에서 두 개의 선별 표지의 측면에 위치한다. 다른 구현예에서, 제2 외인성 핵산은 제6 RRS(즉, 추가적인 RRS)를 제2 유전자와 내의 제4 RRS와 제5 RRS 사이에 포함할 수 있는데, 제4 및 제5 RRS는 발현 카세트의 5' 및 3' 말단에서 두 개의 선별 표지의 측면에 위치한다. RRS들 사이에 포함된 4개의 선별 표지 유전자는 서로 상이하다.

[0024] 많은 구현예에서, 본원에서 제공된 세포는 CHO 세포주의 세포이다.

[0025] 다양한 구현예에서, 사용된 2개의 강화된 발현 유전자좌는 서열번호 1과 적어도 90% 동일한 뉴클레오티드 서열을 포함하는 유전자좌, 서열번호 2와 적어도 90% 동일한 뉴클레오티드 서열을 포함하는 유전자좌, 및 서열번호 3과 적어도 90% 동일한 뉴클레오티드 서열을 포함하는 유전자좌로 이루어지는 군으로부터 선택된다.

[0026] 추가적인 양태에서, 벡터 세트가 세포 내에서 이중 특이적 항원 결합 단백질의 통합 및 발현을 위해 제공된다.

[0027] 일부 구현예에서, 벡터 세트는: 5' 내지 3', 제1 RRS, 제1 HCF를 암호화하는 뉴클레오티드 서열을 함유하는 제1 핵산, 및 제2 RRS를 함유하는 제1 벡터; 5' 내지 3', 제3 RRS, 제2 HCF를 암호화하는 뉴클레오티드 서열을 함유하는 제2 핵산, 및 제4 RRS를 함유하는 제2 벡터; 및 제1 벡터 내의 제1 핵산 내에 있거나 제1 및 제2 벡터와 상이한 제3 벡터 내에 있는 제1 LCF를 암호화하는 뉴클레오티드 서열을 포함하되, 제1, 제2, 제3 및 제4 RRS는 상이하고, 이중 특이적 항원 결합 단백질은 제1 HCF, 제2 HCF 및 제1 LCF를 포함하며, 제1 및 제2 HCF는 상이하다.

[0028] 일부 구현예에서, 제1 LCF를 암호화하는 뉴클레오티드 서열은 제1 벡터 내의 제1 핵산 내에 있다. 일부 구현예에서, 제1 핵산은 제1 선별 표지 유전자를 추가로 포함한다.

[0029] 일부 구현예에서, 제1 LCF를 암호화하는 뉴클레오티드 서열이 제3 벡터 내에 제공되고 5' RRS 및 3' RRS가 그 측면에 위치하되, (i) 3' RRS는 제1 RRS와 동일하고 5' RRS는 제1 및 제2 RRS와 상이하거나, 대안적으로 (ii) 5' RRS는 제2 RRS와 동일하고 3' RRS는 제1 및 제2 RRS와 상이하다. 일부 구현예에서, 제1 및 제3 벡터에 의해 공유된 공통 RRS가 분할된 선별 표지 유전자 포맷(또는 분할 인트론 포맷)으로 제공되도록, 예를 들어, 제1 및 제3 벡터 중 일 벡터 상의 선별 표지 유전자의 5' 부위의 3' 말단에 위치되고, 타 벡터 상의 선별 표지 유전자의 나머지 3' 부위의 5' 단부에 위치되도록 벡터를 설계할 수 있다.

[0030] 일부 구현예에서, 벡터 세트는, 제2 벡터 내의 제2 핵산 내에 제공되거나, 제1, 제2 및 제3 벡터와 별개인 제4 벡터 내에 있는 제2 LCF를 암호화하는 뉴클레오티드 서열을 추가로 포함한다.

[0031] 일부 구현예에서, 제1 및 제2 LCF는 동일하다.

[0032] 일부 구현예에서, 제1 LCF를 암호화하는 뉴클레오티드 서열은 제1 벡터 내의 제1 핵산 내에 포함되고, 제2 VL을 암호화하는 뉴클레오티드 서열은 제4 벡터 상에 제공된다. 일부 구현예에서, 제4 벡터 상의 제2 LCF를 암호화하는 뉴클레오티드 서열의 측면에 5' RRS 및 3' RRS가 위치하되, (i) 3' RRS는 제1 RRS와 동일하고 5' RRS는 제3 및 제4 RRS와 상이하거나, (ii) 5' RRS는 제4 RRS와 동일하고 3' RRS는 제3 및 제4 RRS와 상이하다. 특정 구현예에서, 제2 및 제4 벡터에 의해 공유된 공통 RRS가 (예를 들어, 인트론을 통해) 분할 표지 포맷으로 제공되도록, 예를 들어, 제2 및 제4 벡터 중 일 벡터 상의 선별 표지 유전자의 5' 부위의 3' 말단에 위치되고, 타 벡터 상의 선별 표지 유전자의 나머지 3' 부위의 5' 단부에 위치되도록 벡터를 설계할 수 있다.

[0033] 일부 구현예에서, 제1 LCF를 암호화하는 뉴클레오티드 서열은 제1 벡터 내의 제1 핵산 내에 있고, 제2 VL을 암호화하는 뉴클레오티드 서열은 제2 벡터 상의 제2 핵산 내에 있다.

[0034] 일부 구현예에서, 제1 LCF를 암호화하는 뉴클레오티드 서열은 제1 벡터 상에 있고, 제2 VL을 암호화하는 뉴클레오티드 서열은 제4 벡터 상에 있다. 일부 구현예에서, 제3 벡터 상의 제1 LCF를 암호화하는 뉴클레오티드 서열은 5' RRS 및 3' RRS가 그 측면에 위치하되, (i) 제3 벡터 상의 3' RRS는 제1 RRS와 동일하고, 제3 벡터 상의 5' RRS는 제1 및 제2 RRS와 상이하거나, (ii) 제3 벡터 상의 5' RRS는 제2 RRS와 동일하고, 제3 벡터 상의 3' RRS는 제1 및 제2 RRS와 상이하며; 제4 벡터 상의 제2 LCF를 암호화하는 뉴클레오티드 서열은 5' RRS 및 3' RRS가 그 측면에 위치하되, (i) 제4 벡터 상의 3' RRS는 제3 RRS와 동일하고, 제4 벡터 상의 5' RRS는 제3 및 제4 RRS와 상이하거나, (ii) 제4 벡터 상의 5' RRS는 제4 RRS와 동일하고, 제4 벡터 상의 3' RRS는 제3 및 제4 RRS

와 상이하다.

- [0035] 본원에 제공된 벡터 세트의 많은 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 CH3 도메인을 암호화할 수 있고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제3 CH3 도메인을 암호화할 수 있다. 일부 구현예에서, 제1 및 제2 CH3 도메인은 적어도 하나의 아미노산이 상이하다. 일부 구현예에서, 제1 및 제2 CH3 도메인을 암호화하는 뉴클레오타이드 서열은 뉴클레오타이드 서열 중 하나가 코돈 변형되었다는 점이 상이하다.
- [0036] 본원에 제공된 벡터 세트의 많은 구현예에서, HCF 또는 LCF를 암호화하는 뉴클레오타이드 서열 각각은 프로모터에 독립적으로 연결된다.
- [0037] 일부 구현예에서, 벡터 세트는 LCF 암호화 벡터나 HCF 암호화 벡터 중 하나에 포함되거나 별도 벡터 내에 제공될 수 있는 RRS중 하나 이상을 인식하는 하나 이상의 재조합 효소를 암호화하는 뉴클레오타이드 서열을 추가로 포함할 수 있다.
- [0038] 또 다른 구현예에서, 세포의 제1 발현 강화 유전자좌 내에 통합시키기 위한 5' 상동성 아암 및 3' 상동성 아암이 측면에 위치한 제1 핵산을 함유하는 제1 벡터, 및 세포의 제2 발현 강화 유전자좌 내에 통합시키기 위한 5' 상동성 아암 및 3' 상동성 아암이 측면에 위치한 제2 핵산을 함유하는 제2 벡터를 포함하는 벡터 세트가 제공되며, 제1 및 제2 핵산은 함께 항원 결합 단백질을 암호화한다.
- [0039] 추가적인 양태에서, 본 개시는, 세포(예: CHO 세포)를 하나 이상의 벡터와 재조합하는 것을 포함하는 시스템으로서, 단일 특이적 단백질 또는 이중 특이적 단백질 중 하나인 항원 결합 단백질을 함께 암호화하는 두 개의 발현 강화 유전자좌 외인성 핵산 내에 통합된 세포를 만드는데 사용될 수 있는 시스템을 제공한다. 상기 시스템은, 예를 들어, 키트의 형태로 제공될 수 있다.
- [0040] 특정 구현예에서, 세포 및 벡터의 세트를 포함하는 시스템이 제공되며, 세포는 그의 게놈의 두 개의 별도 발현 강화된 유전자좌 내에서 통합된 RRS 세트로서, 서로 상이하고 벡터 세트 내의 관심 유전자와의 재조합 교환을 위한 하나 이상의 외인성 핵산, 예컨대 선별 표지 사이에서 이격된 RRS 세트를 포함하며, 벡터 세트 내의 RRS는 세포 내의 RRS와 동일한 배치를 포함한다.
- [0041] 일부 구현예에서, 세포 및 벡터 세트를 포함하는 시스템으로서, 상기 세포는, 제1 발현 강화 유전자좌 내에 통합된 제1 RRS, 제1 외인성 핵산, 및 제2 RRS(5'에서 3'까지), 및 제2 발현 강화 유전자좌 내에 통합된 제3 RRS, 제2 외인성 핵산, 및 제4 RRS(5'에서 3'까지)를 포함하며, 제1 및 제2 발현 강화된 유전자좌는 상이하고; 벡터 세트는, (i) 제1 벡터 5' RRS, 제1 핵산, 및 제1 벡터 3' RRS(5'에서 3'까지)를 함유하되, 제1 벡터 5' RRS와 제3' RRS는 상이한 제1 벡터, (ii) 제2 벡터 5' RRS, 제2 핵산, 및 제2 벡터 3' RRS(5'에서 3'까지)를 함유하되, 제2 벡터 5' RRS와 3' RRS는 상이한 제2 벡터, 및 (iii) 제1 HCF를 암호화하는 뉴클레오타이드 서열 및 제1 LCF를 암호화하는 뉴클레오타이드 서열로서, 중쇄를 암호화하는 2개의 뉴클레오타이드 서열 중 하나는 제1 핵산 내에 있고 다른 하나의 뉴클레오타이드 서열은 제2 핵산 내에 있는 뉴클레오타이드 서열을 포함하고, 제1 HCF 및 제1 LCF는 항원 결합 도메인의 영역이며, 세포 내에 벡터의 도입 시, 벡터 내의 제1 및 제2 핵산은 RRS에 의해 매개된 재조합을 통해 제1 발현 강화 유전자좌 및 제2 발현 강화 유전자좌 내에 각각 통합되는 시스템이 제공된다.
- [0042] 일부 구현예에서, 항원 결합 단백질은 단일 특이적 항원 결합 단백질이다.
- [0043] 일부 구현예에서, 제1 RRS와 제3 RRS가 동일하고, 제2 RRS와 제4 RRS가 동일하다. 특정 구현예에서, 제1 추가적인 RRS는 제1 유전자좌에서 제1 및 제2 RRS 사이에 존재한다. 일부 구현예에서, 제1 벡터 5' RRS는 제1 및 제3 RRS와 동일하고; 제1 벡터 3' RRS, 제2 벡터 5' RRS, 및 제1 추가 RRS가 동일하고; 제2 벡터 3' RRS는 제2 및 제4 RRS와 동일하다. 일부 구현예에서, LCF를 암호화하는 뉴클레오타이드 서열은 제1 벡터 내에 있고, HCF를 암호화하는 뉴클레오타이드 서열은 제2 벡터 내에 있다. 일부 구현예에서, 제1 벡터 3' RRS는 선별 표지 유전자의 5' 부분의 3' 말단에 위치하고, 제2 벡터 5' RRS는 선별 표지 유전자의 나머지 3' 부분의 5' 말단에 위치한다. 다른 구현예에서, 제1 벡터 5' RRS는 제1 RRS와 동일하고, 제1 벡터 3' RRS는 제2 RRS와 동일하며; 제2 벡터 5' RRS는 제3 RRS와 동일하고, 제2 벡터 3' RRS는 제4 RRS와 동일하다.
- [0044] 다양한 구현예에서, 항원 결합 단백질은 이중 특이적 항원 결합 단백질이다.
- [0045] 일부 구현예에서, 시스템 내의 벡터 세트는 제1 HCF와 상이한 제2 HCF를 암호화하는 뉴클레오타이드 서열을 추가로 포함한다.
- [0046] 일부 구현예에서, 제1 LCF를 암호화하는 뉴클레오타이드 서열과 제2 HCF를 암호화하는 뉴클레오타이드 서열 둘 다는 제1 벡터 내의 제1 핵산에 포함되고, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제2 벡터 내에 있다. 일부 구

현예에서, 제1 벡터 5' RRS는 제1 RRS와 동일하고, 제1 벡터 3' RRS는 제2 RRS와 동일하고, 제2 벡터 5' RRS는 제3 RRS와 동일하고, 제2 벡터 3' RRS는 제4 RRS와 동일하다.

[0047] 일부 구현예에서, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제3 벡터 5' RRS 및 제3 벡터 3' RRS가 측면에 위치한 별도의 제3 벡터 상에 있다. 일부 구현예에서, 제1 LC를 암호화하는 뉴클레오타이드 서열은 제1 벡터 내에 있고, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제2 벡터 내에 있고, 제1 벡터 5' RRS는 제1 RRS와 동일하고, 제1 벡터 3' RRS는 제2 벡터 5' RRS 및 제1 추가 RRS와 동일하고, 제2 벡터 3' RRS는 제2 RRS와 동일하고, 제3 벡터 5' RRS는 제3 RRS와 동일하고, 제3 벡터 3' RRS는 제4 RRS와 동일하며, 제1 추가 RRS는 제1 및 제2 RRS 사이의 제1 유전자좌에 포함된다. 일부 구현예에서, 벡터는 분할 표지 포맷으로 공통 RRS를 제공하도록, 예를 들어, 제1 벡터 3' RRS가 제1 벡터에 포함된 선별 표지 유전자의 5' 부분의 3' 말단에 위치되고, 제2 벡터 5' RRS가 제2 벡터에 포함된 나머지 선별 표지 유전자의 5' 말단에 위치하도록, 설계될 수 있다.

[0048] 일부 구현예에서, 시스템의 벡터 세트는 제1 LCF와 동일하거나 상이할 수 있는 제2 LCF를 암호화하는 뉴클레오타이드 서열을 추가로 포함한다.

[0049] 일부 구현예에서, 제2 LCF를 암호화하는 뉴클레오타이드 서열은 제2 벡터의 제2 핵산 내에 있되, 제1 벡터 5' RRS는 제1 RRS와 동일하고, 제1 벡터 3' RRS는 제2 RRS와 동일하고, 제2 벡터 5' RRS는 제3 RRS와 동일하며, 제2 벡터 3' RRS는 제4 RRS와 동일하다.

[0050] 일부 구현예에서, 제2 LCF를 암호화하는 뉴클레오타이드 서열은 제3 벡터 5' RRS 및 제3 벡터 3' RRS가 측면에 위치한 별도의 제3 벡터 내에 있다. 일부 구현예에서, 제1 벡터 5' RRS 및 3' RRS는 제1 유전자좌 내의 제1 RRS 및 제2 RRS와 동일하고, 제3 벡터 5' RRS는 제3 RRS와 동일하고, 제3 벡터 3' RRS는 제2 벡터 5' RRS 및 제2 유전자좌 내에서 제3 RRS와 제4 RRS 사이에 존재하는 추가 RRS와 동일하며, 제2 벡터 3' RRS는 제4 RRS와 동일하다. 일부 구현예에서, 공통 RRS는 분할 표지 포맷으로, 예를 들어, 제3 벡터 3' RRS가 제3 벡터에 포함된 선별 표지 유전자의 5' 부분의 3' 말단에 위치되고, 제2 벡터 5' RRS가 제2 벡터에 포함된 선별 표지 유전자의 나머지 3' 부분의 5' 말단에 위치하도록, 설계될 수 있다.

[0051] 본원에 제공된 시스템의 많은 구현예에서, HCF 또는 LCF를 암호화하는 뉴클레오타이드 서열은 불변 영역으로부터 아미노산을 암호화할 수 있다. 일부 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 CH3 도메인을 암호화할 수 있고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 CH3 도메인을 암호화할 수 있다. 일부 구현예에서, 제1 및 제2 CH3 도메인은 적어도 하나의 아미노산이 상이하다. 일부 구현예에서, 제1 및 제2 CH3 도메인을 암호화하는 뉴클레오타이드 서열은 뉴클레오타이드 서열 중 하나가 코돈 변형되었다는 점이 상이하다.

[0052] 본원에 제공된 시스템의 많은 구현예에서, HCF 또는 LCF를 암호화하는 뉴클레오타이드 서열 각각은 프로모터에 독립적으로 연결된다.

[0053] 일부 구현예에서, 시스템의 벡터 세트는 LCF 암호화 벡터나 HCF 암호화 벡터 중 하나에 포함되거나 별도의 벡터로 제공될 수 있는 RRS중 하나 이상을 인식하는 하나 이상의 재조합 효소를 암호화하는 뉴클레오타이드 서열을 추가로 포함할 수 있다.

[0054] 다양한 구현예에서, 본원에 제공된 시스템 내의 세포는 CHO 세포이다.

[0055] 다양한 구현예에서, 두 개의 발현 강화된 유전자좌는 서열번호 1의 뉴클레오타이드 서열을 포함하는 유전자좌, 서열번호 2의 뉴클레오타이드 서열을 포함하는 유전자좌, 및 서열번호 3의 뉴클레오타이드 서열을 포함하는 유전자좌로 구성되는 군으로부터 선택된다.

[0056] 또 다른 양태에서, 본 개시는 또한 이중 특이적 항원 결합 단백질을 만드는 방법을 제공한다. 일 구현예에서, 상기 방법은 본원에 개시된 시스템을 사용하며, 시스템의 벡터를 시스템의 세포 내에 형질감염(transfection)에 의해 도입한다. 외인성 핵산이 2개의 발현 강화된 유전자좌 내에 RMCE를 통해 적절하게 통합된 형질감염된 세포는 스크리닝되고 동정될 수 있다. HCF 함유 폴리펩티드 및 LCF 함유 폴리펩티드는 통합된 핵산으로부터 발현될 수 있고, 항원 결합 관심 단백질은 동정된 형질감염 세포로부터 수득될 수 있고, 공지된 방법을 사용해 정제될 수 있다.

[0057] 또 다른 구현예에서, 상기 방법은, 항원 결합 단백질을 함께 암호화하고 세포로부터 상기 항원 결합 단백질을 발현시키는 2개의 발현 강화 유전자좌에 통합된 외인성 핵산을 함유하는, 상기 본원에 기술된 세포를 단순히 사용한다.

도면의 간단한 설명

[0058]

도 1. 통상적인 단일 특이적 항체에 대해, 다중 발현 강화 유전자좌와 비교하여 하나의 발현 강화 유전자좌 내에서의 통합을 위한 예시적인 항체 클로닝 전략. 2개의 벡터, 즉 경쇄와 같은 항체 사슬 1(AbC1)을 암호화하는 핵산을 가지는 제1 벡터와, 중쇄와 같은 항체 사슬 2(AbC2)를 암호화하는 핵산 및 숙주 세포의 표적화 유전자좌 내에 통합된 표지와 상이한 선별 표지를 가지는 제2 벡터를 숙주 세포 내로 형질감염시켰다. (5'에서 3'까지) 벡터 작제물의 RRS1, 중간 RRS3 및 RRS2 부위는 숙주 세포의 유전자좌 내에서 선별 표지의 측면에 위치하는 RRS 부위와 일치한다. 숙주 세포 내로 형질감염된 추가 벡터가 재조합 효소를 암호화한다. 제2 벡터의 선별 표지가 항생제 내성 유전자인 경우, 2개의 벡터가 결합되어 표지를 발현시키도록 조작되므로, 양성 재조합 클론이 항생제 내에서의 성장을 위해 선택된다. 대안적으로, 형광 표지가 형광 활성 세포 분류(FACS) 분석에 의해 양성 클론 선택을 가능하게 한다. 동일한 벡터가 EESYR® 유전자좌(1번 유전자좌)와 같은 단일 유전자좌에서의 부위 특이적 통합에 사용될 수 있다.

도 2. 2개의 유전자좌 대신에 1개의 발현 강화 유전자좌 내에서의 통합을 위한 2개의 벡터 클로닝 전략을 사용하여, 항체 A 및 항체 B를 도 1에 도시된 바와 같이 유전자좌 내로 클로닝하였다. 항체 발현 세포를 단리하고 12일간 유가배양 배지에서 배양한 뒤 수확하고, 단백질 A 센서를 사용해 옥텟(Octet) 역가 검정을 실시하였다. 세포는 또한 동질유전자이며 안정적인 것으로 관찰되었다. 총 역가는 2개 부위 통합 방법을 사용함으로써 0.5 내지 0.9배 증가하는 것으로 관찰되었다.

도 3. 3개의 항체 사슬에 의해 암호화된 이중 특이적 항체에 대해, 2개의 분리된 발현 강화 유전자좌 내에서의 통합과 비교하여 하나의 발현 강화 유전자좌 내에서의 통합을 위한 예시적인 항체 클로닝 전략. 본 이중 특이적 전략에는 3개의 벡터, 즉 예를 들어, EESYR® (서열번호 1; 1번 유전자좌) 내로의 통합을 위해 RRS1 및 RRS3이 측면에 위치하는, 공통 경쇄와 같은 항체 사슬 1(AbC1)을 암호화하는 핵산을 가지는 제1 벡터; 서열번호 2 유전자좌 내로의 통합을 위해 RRS4 및 RRS6이 측면에 위치하는 상류 선별 표지를 갖는, 중쇄와 같은 항체 사슬 2(AbC2)을 가지는 제2 벡터; 및 제2 벡터 내의 선별 표지와 상이한 선별 표지에 연결되고, 벡터 카세트 내의 5' RRS4 및 3' RRS5가 측면에 위치한(5' RRS와 3' RRS는 서열번호 2를 포함하는 유전자좌에서의 숙주 세포 내의 RRS 부위와 일치함), 제2 (상이한) 중쇄와 같은 항체 사슬 3(AbC3)에 연결된 제2 카피 AbC1을 암호화하는 핵산을 가지는 추가적인 제3 벡터를 사용하였다. 도시된 바와 같은 2벡터 시스템이 EESYR® 유전자좌(서열번호 1; 1번 유전자좌를 포함함)와 같은 단일 유전자좌에서의 부위 특이적 통합에 사용될 수 있다. 각각의 생산 세포주의 역가를 분석하였다(도 5 참조).

도 4. 3개 또는 4개의 항체 사슬에 의해 암호화된 이중 특이적 항체에 대해, 2개의 분리된 발현 강화 유전자좌 내에서의 통합과 비교하여 하나의 발현 강화 유전자좌 내에서의 통합을 위한 예시적인 항체 클로닝 전략. 본 이중 특이적 전략에는 4개의 벡터, 즉 EESYR® (서열번호 1; 1번 유전자좌) 내로의 통합을 위해 RRS1 및 RRS3이 측면에 위치하는, 제1 경쇄와 같은 항체 사슬 1(AbC1)을 암호화하는 핵산을 가지는 제1 벡터; 또한 EESYR® 유전자좌(서열번호 1; 유전자좌 1을 포함함) 내로의 통합을 위해 5' RRS3 및 3' RRS2가 측면에 위치하는 상류 선별 표지를 갖는, 중쇄와 같은 항체 사슬 2(AbC2)를 가지는 제2 벡터; 제2 벡터 내의 선별 표지와 상이한 선별 표지를 암호화하고, 벡터 카세트 내의 5' RRS6 및 3' RRS5가 측면에 위치한(5' RRS와 3' RRS는 서열번호 2, 2번 유전자좌를 포함하는 유전자좌에서의 숙주 세포 내의 RRS 부위와 일치함), 제2 (상이한) 중쇄와 같은 항체 사슬 3(AbC3)에 연결된 핵산을 가지는 제3 벡터; 및 제2 경쇄(제1 경쇄와 동일하거나 상이할 수 있음), 예를 들어 항체 사슬 1(AbC1)을 암호화하는 핵산을 가지는 추가적인 제4 벡터를 사용하였다. 도시된 바와 같은 4벡터 시스템이 EESYR® 유전자좌(서열번호 1; 1번 유전자좌를 포함함) 및 서열번호 2 또는 3을 포함하는 유전자좌와 같은 2개의 유전자좌에서의 부위 특이적 통합에 사용될 수 있다. 4벡터 시스템은 하나의 유전자좌(EESYR® 유전자좌; 1번 유전자좌)에서의 2벡터 시스템 통합과 비교된다.

도 5. 2개의 유전자 대신에, 하나의 발현 강화 유전자좌 내에서의 통합을 위한 2벡터 클로닝 전략을 사용하여, 이중 특이적 항체 C 및 이중 특이적 항체 D를 클로닝하고, 세포를 단리하고, 12일간의 유가배양 배지에서 배양한 뒤 수확하고, 고정화 항Fc 및 제2 항Fc*(변형된 Fc 검출 항체, 2014년5월15일 공개된 US 2014-0134719 A1 참조)를 사용하여 옥텟 역가 검정을 수행하였다. 세포는 또한 동질유전자이며 안정적인 것으로 관찰되었다. 총 이중 특이적 역가(이중 이량체 형성)는 하나의 통합 부위 내의 이중 특이적 항체(이중 이량체)의 발현과 비교하면, 2개 부위 통합 방법을 사용하여 1.75배 내지 2배 증가하는 것으로 관찰되었다.

도 6a 및 6b. 이중 특이적 항체 E(Ab E), 이중 특이적 항체 F(Ab F), 이중 특이적 항체 G(Ab G), 및 이중 특이

적 항체 H(Ab H)를 RSX 또는 RSX^{2BP}에서 클로닝하였다. 각각의 이중 특이적 항체 발현 세포는 하나의 발현 강화 유전자좌(RSX) 또는 2개의 발현 강화 유전자좌(RSX^{2BP}) 중 하나에서 (공통) 경쇄 뉴클레오타이드, 중쇄 뉴클레오타이드(야생형 Fc) 및 변형된 중쇄 뉴클레오타이드(Fc*)를 포함한다. 세포를 단리하고, 바이오리액터 내의 유가배양 배지에서 13일간 배양한 뒤, 수확하고 HPLC 용출 방법을 수행하여 총 항체 및 이중 특이적 항체 역가를 결정하였다(도 6a). 총 항체 역가 당 이중 특이적 항체 종의 역가(동종 이량체 종으로부터 정제됨)의 비율을 총 Ab 백분율로서 결정하였다(도 6b).

도 7. 단일 특이적 항체 J 및 K(각각 Ab J 및 Ab K)를 RSX 또는 RSX²에서 클로닝하였다. 세포를 단리하고, 바이오리액터 내의 유가배양 배지에서 13일간 배양한 뒤, 수확하고 HPLC 방법을 수행하여 총 IgG 역가를 결정하였다.

발명을 실시하기 위한 구체적인 내용

[0059] 정의

용어 "항체"는 본원에서 사용되는 바와 같이, 4개의 폴리펩티드 사슬, 즉 이황화 결합에 의해 상호 연결된 2개의 중(H)쇄 및 2개의 경(L)쇄를 포함하는 면역 글로불린 분자를 포함한다. 각각의 중쇄는 중쇄 가변 영역(본원에서 HCVR 또는 VH로 약칭됨)과 중쇄 불변 영역을 포함한다. 중쇄 불변 영역은 3개의 도메인 CH1, CH2 및 CH3을 포함한다. 각각의 경쇄는 경쇄 가변 영역(본원에서 LCVR 또는 VL로 약칭됨)과 경쇄 불변 영역을 포함한다. 경쇄 불변 영역은 하나의 도메인, CL을 포함한다. VH와 VL 영역은 추가적으로, 더 보존적이면서 프레임워크 영역(FR)인 영역들이 중간에 끼어 있는, 상보성 결정 영역(CDR)이라 지칭되는 초가변성 영역으로 다시 나누어질 수 있다. 각각의 VH 및 VL은 아미노 말단으로부터 카르복시 말단 쪽으로 다음 순서로 배열된 3개의 CDR과 4개의 FR을 포함한다: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4 (중쇄 CDR은 HCDR1, HCDR2 및 HCDR3로서 약칭될 수 있고; 경쇄 CDR은 LCDR1, LCDR2 및 LCDR3로서 약칭될 수 있음).

"항원 결합 단백질"이라는 문구는, 적어도 하나의 CDR을 가지며 항원을 선택적으로 인식할 수 있는, 즉 항원을 적어도 마이크로몰 범위에 있는 μM 과 결합시킬 수 있는 단백질을 포함한다. 치료용 항원 결합 단백질(예: 치료 항체)는 나노몰 또는 피코몰 범위에 있는 μM 를 빈번하게 필요로 한다. 일반적으로, 항원 결합 단백질은 2개 이상의 CDR, 예를 들어 2, 3, 4, 5 또는 6개의 CDR을 포함한다. 예시적인 항원 결합 단백질은 항체, 항체의 중쇄 및 경쇄 가변 영역을 함유하는 폴리펩티드와 같은 항체의 항원 결합 단편(예: Fab 단편, F(ab')₂ 단편), 및 항체의 중쇄 및 경쇄 가변 영역을 함유하고 중쇄 및/또는 경쇄의 불변 영역(예컨대, 하나 이상의 불변 영역, 즉, CL, CH1, 힌지, CH2, 및 CH3 도메인 중 하나 이상)으로부터 추가 아미노산을 함유하는 단백질을 포함한다.

"이중 특이적 항원 결합 단백질"이라는 문구는 2개의 상이한 분자(예: 항원) 또는 동일한 분자(예: 동일한 항원) 상에서 2개 이상의 에피토프에 선택적으로 결합할 수 있거나, 상기 2개 이상의 에피토프에 대해 상이한 특이성을 갖는 항원 결합 단백질을 포함한다. 이러한 단백질의 항원 결합 부분, 또는 단편 항원 결합(Fab) 부분은 특정 항원에 대한 특이성을 나타내고, 일반적으로 면역 글로불린의 중쇄 가변 영역 및 경쇄 가변 영역으로 구성된다. 일부 환경에서, 중쇄 가변 영역 및 경쇄 가변 영역은 동족 쌍이 아닐 수 있는데, 달리 말해, 상이한 결합 특이성을 가질 수 있다.

예시적 특이적 항원 결합 단백질은 "이중 특이적 항체"이며, 2개 이상의 에피토프에 선택적으로 결합할 수 있는 항체를 포함한다. 이중 특이적 항체는 일반적으로 2개의 상이한 중쇄를 포함하며, 각각의 중쇄는 2개의 상이한 분자 (예, 항원) 또는 동일한 분자 (예, 동일한 항원) 상에서 상이한 에피토프에 특이적으로 결합한다. 이중 특이적 항체가 2개의 상이한 에피토프 (제1 에피토프 및 제2 에피토프)에 선택적으로 결합할 수 있는 경우, 제1 에피토프에 대한 제1 중쇄의 가변 영역의 친화성은 일반적으로 제2 에피토프에 대한 제1 중쇄의 가변 영역의 친화성보다 적어도 10배 내지 100배, 또는 1,000배 내지 10,000배 낮아지거나 그 반대가 될 것이다. 이중 특이적 항체와 같은 이중 특이적 항원 결합 단백질은 동일한 항원의 상이한 에피토프를 인식하는 중쇄의 가변 영역을 포함할 수 있다. 일반적인 이중 특이적 항체는 2개의 중쇄를 갖는데, 이들 각각은 (N 말단에서 C 말단까지) CH1 도메인, 힌지, CH2 도메인 및 CH3 도메인이 뒤따르는 3개의 중쇄 CDR, 및 면역 글로불린 경쇄를 갖는데, 상기 면역 글로불린 경쇄는 항체 결합 특이성을 부여하지 않지만 각각의 중쇄와 결합할 수 있거나, 각각의 중쇄와 결합하면서 중쇄 항원 결합 영역에 의해 결합된 에피토프 중 하나 이상과 결합할 수 있거나, 각각의 중쇄와 결합할 수 있고 중쇄 중 하나 또는 둘 모두를 하나 또는 둘 모두의 에피토프에 결합시킬 수 있다. 일 구현예에서,

Fc 도메인은 적어도 CH2 및 CH3를 포함한다. Fc 도메인은 힌지, CH2 도메인 및 CH3 도메인을 포함할 수 있다.

[0064] 구현된 하나의 이중 특이적 포맷은 제1 중쇄(HC), 변형된 CH3을 갖는 제2 중쇄(HC*), 및 공통 경쇄(LC)(동일한 경쇄의 2개의 카피)를 포함한다. 또 다른 구현예에는 제1 중쇄(HC), 공통 LC 및 (제2 HC가 ScFv의 N 말단에 융합된) HC-ScFv 융합 폴리펩티드가 포함된다. 또 다른 구현예에는 제1 HC, 동족 LC, 및 (제2 HC가 ScFv의 N 말단에 융합된) HC-ScFv 융합 폴리펩티드가 포함된다. 또 다른 구현예에는 제1 중쇄(HC), LC 및 Fc 도메인이 포함된다. 또 다른 구현예에는 제1 HC, LC, 및 (Fc가 ScFv의 C 말단에 융합된) ScFv-Fc 융합 폴리펩티드가 포함된다. 또 다른 구현예에는 제1 HC, 공통 LC, 및 (Fc가 ScFv의 N 말단에 융합된) Fc-ScFv 융합 폴리펩티드가 포함된다. 또 다른 구현예에는 제1 HC, LC, 및 (제2 HC가 ScFv의 C 말단에 융합된) ScFv-HC가 포함된다.

[0065] 특정 구현예에서, 하나의 중쇄(HC)는 원시 서열 또는 "야생형" 서열일 수 있고, 제2 중쇄는 Fc 도메인에서 변형될 수 있다. 다른 구현예에서, 하나의 중쇄(HC)는 원시 서열 또는 "야생형" 서열일 수 있고, 제2 중쇄는 코돈 변형될 수 있다.

[0066] 용어 "세포"는, 재조합 핵산 서열을 발현하는데 적합하고, 외인성 핵산을 안정적으로 통합시키고 발현을 강화할 수 있는 유전자좌를 갖는 임의의 세포를 포함한다. 세포는 비인간 동물 세포와 같은 포유류 세포, 인간 세포, 또는 예를 들어 하이브리도마 또는 쿼드라마와 같은 세포 융합체를 포함한다. 일부 구현예에서, 세포는 인간, 원숭이, 유인원, 햄스터, 랫트, 또는 마우스 세포이다. 일부 구현예에서, 세포는 다음의 세포로부터 선택된 포유류 세포이다: CHO(예를 들어, CHOK1, DXB-11 CHO, Veggie-CHO), COS(예를 들어, COS-7), 망막 세포, Vero, CV1, 신장(예를 들어, HEK293, 293 EBNA, MSR 293, MDCK, HaK, BHK), HeLa, HepG2, WI38, MRC 5, Colo205, HB 8065, HL-60, (예를 들어, BHK21), Jurkat, Daudi, A431(상피), CV-1, U937, 3T3, L 세포, C127세포, SP2/0, NS-0, MMT 060562, Sertoli 세포, BRL 3A 세포, HT1080 세포, 골수종 세포, 종양 세포, 및 전술된 세포로부터 유래한 세포주. 일부 구현예에서, 세포는 하나 이상의 바이러스 유전자, 예를 들어, 바이러스 유전자를 발현하는 망막 세포(예를 들어, PER.C6TM 세포)를 포함한다.

[0067] "세포 밀도"는, 예를 들어, mL 당 총 (생존 및 사멸) 세포 수로서의 시료의 부피 당 세포 수를 지칭한다. 세포 수는 수동으로 또는 자동으로, 예컨대 유동 세포 계측기로 계수할 수 있다. 자동화된 세포 계수기는 예를 들어 표준 트립판 블루 흡수 기술을 사용하여 생존 세포 또는 사멸 세포 또는 생존/사멸 세포의 수를 계수하기에 적합하도록 구성되었다. "생존 세포 밀도" 또는 "생존 세포 농도"라는 구문은 시료의 부피 당 생존 세포의 수를 지칭한다("생존 세포 계수"로서도 지칭됨). 세포 밀도를 결정하기 위해 임의의 수의 잘 알려진 수동 또는 자동화 기술이 사용될 수 있다. 커패시턴스 또는 광학 밀도가 부피 당 세포 수와 상관된, 배양물의 온라인 바이오매스 측정치가 측정될 수 있다. 생산 배양물과 같은 세포 배양물에서의 최종 세포 밀도는 출발 세포주에 따라, 예를 들어, 약 1.0 내지 10×10^6 세포/mL의 범위 내에서 달라진다. 일부 구현예에서, 최종 세포 밀도는 생산 세포 배양물로부터 관심 단백질을 수확하기 전에 1.0 내지 10×10^6 세포/mL에 도달한다. 다른 구현예에서, 최종 세포 밀도는 5.0×10^6 세포/mL 초과, 6×10^6 세포/mL 초과, 7×10^6 세포/mL 초과, 8×10^6 세포/mL 초과, 9×10^6 세포/mL 초과, 또는 10×10^6 세포/mL 초과에 도달한다.

[0068] 용어 "코돈 변형"은 코돈에 의해 암호화된 아미노산을 변화시키지 않고 단백질 코딩 서열이 하나 이상의 뉴클레오타이드, 즉 하나 이상의 코돈에서 변형되어 뉴클레오타이드 서열의 코돈 변형된 버전이 생성되는 것을 의미한다. 뉴클레오타이드 서열의 코돈 변형은 핵산에 기초한 검정(예를 들어, 특히 하이브리드화에 기초한 검정, PCR 등)에서 뉴클레오타이드 서열을 그의 코돈 변형 버전과 구별하기에 편리한 기초를 제공할 수 있다. 일부 실시예에서, 뉴클레오타이드 서열의 코돈은 숙주 세포에서 암호화된 단백질의 개선되거나 최적화된 발현을 제공하도록 당업계에 잘 알려진 코돈 최적화 기술을 사용하여 변형된다(Gustafsson, C. 외의 2004, *Trends in Biotechnology*, 22:346-353; Chung, B.K.-S. 외의 2013, *Journal of Biotechnology*, 167:326-333; Gustafsson, C. 외의 2012, *Protein Expr Purif*, 83(1): 37-46). 이러한 기술을 사용하는 서열 설계 소프트웨어 툴 또한 당업계에 잘 알려져 있는데, 이에는, Codon Optimizer (Fuglsang A. 2003, *Protein Expr Purif*, 31:247-249), Gene Designer (Villalobos A, 외의 2006, *BMC Bioinforma*, 7:285), 및 OPTIMIZER (Puigbo P, 외의 2007, *Nucleic Acids Research*, 35:W126-W131)가 특히 포함되지만 이들로 한정되지 않는다.

[0069] 문구 "상보성 결정 영역", 또는 용어 "CDR"은 보통 (즉, 야생형 동물에서) 면역 글로불린 분자(예: 항체 또는 T 세포 수용체)의 경쇄 또는 중쇄 가변 영역의 2개의 프레임워크 영역 사이에 나타나는, 유기체의 면역 글로불린 유전자의 핵산 서열에 의해 암호화된 아미노산 서열을 포함한다. CDR은 예를 들어, 생식계열 서열 또는 재배열

되거나 재배열되지 않은 서열에 의해 암호화될 수 있으며, 예를 들어, 미처리 B 세포 또는 성숙한 B 세포 또는 T 세포에 의해 암호화될 수 있다. 일부 경우에(예: CDR3의 경우), CDR은, (예컨대, 재배열되지 않은 핵산 서열에서는) 서로 인접하지 않지만 B 세포 핵산 서열 내에서는 인접하는 2개 이상의 서열(예: 생식계열 서열)에 의해, 예를 들어 서열의 접합 또는 연결(예를 들어, 중쇄 CDR3을 형성하기 위한 V-D-J 재조합)의 결과로서 암호화될 수 있다.

[0070] 용어 "발현 강화 유전자좌"는 세포의 게놈 내의 유전자좌로서, 적합한 유전자 또는 작제물이 서열/서열들 내에 또는 이에 가깝게 외인성으로 첨가(즉, 통합)되거나, 서열/서열들에 "작동 가능하게 연결"될 때, 게놈 내의 다른 영역 또는 서열과 비교하여 더 높은 레벨의 발현을 나타내는 서열/서열들을 함유하는 유전자좌를 지칭한다.

[0071] "강화된"이라는 용어가 강화된 발현을 설명하도록 사용되는 경우, 예를 들어, 동일 발현 작제물의 단일 카피의 무작위 인티그란트(integrants) 풀과 비교하여, 외인성 서열을 게놈 내에 무작위 통합하거나 상이한 유전자좌에서 통합하는 것에 의해 일반적으로 관찰되는 것에 비해 발현이 적어도 1.5배 내지 적어도 3배 강화된 것을 포함한다. 실질적으로 동일한 조건에서, 본 발명의 서열의 부재에서 동일한 유전자의 발현 수준에 비해, 예컨대 동일한 종의 게놈 내의 또 다른 유전자좌에서 통합시키는 경우에 비해, 본 발명의 서열을 사용할 때 수 배의 발현 강화가 관찰되었다. 강화된 재조합 효율은 (예를 들어, 재조합 효소-인식 부위("RRS")를 이용하여) 재조합할 수 있는 유전자좌의 능력 강화를 포함한다. 강화는 재조합 효소 인식 부위 등을 사용하지 않은 무작위 재조합(일반적으로 0.1%임)에 대한 재조합의 효율을 지칭한다. 바람직한 강화된 재조합 효율은 무작위에 비해 약 10배이거나 약 1%이다. 명시되지 않는 한, 청구된 발명은 특정한 재조합 효율로 제한되지 않는다. 강화된 발현 유전자좌는 일반적으로 숙주 세포에 의한 관심 단백질의 높은 생산성을 지지한다. 따라서, 강화된 발현은 단순히 배양된 세포의 높은 카피 수에 의해 높은 역가를 달성하기 보다는, 세포 당 많은 관심 단백질을 생산하는 것(상승된 그램 당 단백질 역가)을 포함한다. 특정 생산성 Qp(pg/세포/일, 즉 pcd)는 지속 가능한 생산성의 척도로 간주된다. 5 pcd보다 크거나, 10 pcd보다 크거나, 15 pcd보다 크거나, 20 pcd보다 크거나, 25 pcd보다 크거나, 심지어 30 pcd보다 큰 Qp를 나타내는 재조합 숙주 세포가 바람직하다. 발현 강화 유전자좌, 또는 "핫스팟(hotspot)" 내에 삽입된 관심 유전자를 갖는 숙주 세포는 높은 특정 생산성을 나타낸다.

[0072] 문구 "외인성으로 추가된 유전자", "외인성으로 추가된 핵산", 또는 간단히 "외인성 핵산"이 관심 유전자좌와 관련하여 사용되는 경우, 이러한 문구는 유전자좌가 자연에서 발견될 때 관심 유전자좌 내에 존재하지 않는 임의의 DNA 서열 또는 유전자를 지칭한다. 예를 들어, CHO 유전자좌(예컨대, 서열번호 1 또는 서열번호 2의 서열을 포함하는 유전자좌) 내의 "외인성 핵산"은 자연에서 특정 CHO 유전자좌 내에서 발견되지 않은 햄스터 유전자(즉, 햄스터 게놈 내의 또 다른 유전자좌의 햄스터 유전자), 임의의 다른 종의 유전자(예컨대, 인간 유전자), 키메라 유전자(예컨대, 인간/마우스), 또는 관심 CHO 유전자좌 내에 존재하는 것으로 자연에서 발견되지 않은 임의의 다른 유전자일 수 있다.

[0073] "중쇄" 또는 "면역 글로불린 중쇄"라는 문구는 임의의 유기체의 면역 글로불린 중쇄 불변 영역 서열을 포함하며, 달리 명시되지 않는 한 중쇄 가변 도메인을 포함한다. 중쇄 가변 도메인은 달리 명시되지 않는 한, 3개의 중쇄 CDR 및 4개의 FR 영역을 포함한다. 일반적인 중쇄는, (N 말단에서 C 말단까지) 가변 도메인에 이어서, CH1 도메인, 힌지, CH2 도메인 및 CH3 도메인을 갖는다. 용어 "중쇄 단편"은 중쇄의 적어도 10, 20, 30, 40, 50, 60, 70, 80, 90 또는 100개 이상의 아미노산으로 이루어진 펩티드를 포함하고, 하나 이상의 CDR, 하나 이상의 FR과 결합된 하나 이상의 CDR, CH1, 힌지, CH2 또는 CH3 중 하나 이상, 가변 영역, 불변 영역, 불변 영역의 단편(예: CH1, CH2, CH3) 또는 이들의 조합을 포함할 수 있다. 예시적인 HCF는 VH를 포함하고, Fc 영역의 전체 또는 일부를 포함한다. 문구 "HCF를 암호화하는 뉴클레오티드 서열"은 HCF로 이루어지는 폴리펩티드를 암호화하는 뉴클레오티드 서열 및 HCF를 함유하는 폴리펩티드(예를 들어, 특정 HCF 외에 추가 아미노산을 함유할 수 있는 폴리펩티드)를 암호화하는 뉴클레오티드 서열을 포함한다. 예를 들어, HCF를 암호화하는 뉴클레오티드 서열은, 특히, VH로 이루어지는 폴리펩티드, CH3에 연결된 VH로 이루어지는 폴리펩티드, 전체 중쇄로 이루어지는 폴리펩티드를 암호화하는 뉴클레오티드 서열을 포함한다.

[0074] 핵산 서열의 맥락에서 "상동 서열"은 참조 핵산 서열과 실질적으로 상동인 서열을 지칭한다. 일부 구현예에서, 2개의 서열의 대응하는 뉴클레오티드의 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 이상이 잔기의 관련 구간에 걸쳐 동일한 경우 2개의 서열은 실질적으로 상동인 것으로 간주된다. 일부 구현예에서, 관련 구간은 완전한 (즉, 전장) 서열이다.

[0075] "경쇄"라는 문구는 임의의 유기체의 면역 글로불린 경쇄 불변 영역 서열을 포함하며, 달리 명시되지 않는 한 인간 카파(kappa) 및 람다(lambda) 경쇄를 포함한다. 경쇄 가변 (VL) 도메인은 달리 명시되지 않는 한, 일반적으

로 3개의 경쇄 CDR 및 4개의 프레임워크(FR) 영역을 포함한다. 일반적으로, 전장 경쇄는 아미노 말단에서 카르복실 말단까지, FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4를 포함하는 VL 도메인, 및 경쇄 불변 도메인을 포함한다. 본 발명과 함께 사용될 수 있는 경쇄는, 예를 들어, 이중 특이적 항체에 의해 선택적으로 결합된 제1 에피토프 또는 제2 에피토프에 선택적으로 결합하지 않는 것들을 포함한다. 적절한 경쇄는 항체의 항원 결합 영역에 의해 결합된 하나 또는 둘 모두의 에피토프에 결합할 수 있거나 이의 결합에 기여할 수 있는 것들을 포함한다. 용어 "경쇄의 단편" 또는 "경쇄 단편"(또는 "LCF")는 경쇄의 적어도 10, 20, 30, 40, 50, 60, 70, 80, 90 또는 100 개 이상의 아미노산으로 이루어진 펩티드를 포함하고, 하나 이상의 CDR, 하나 이상의 FR과 결합된 하나 이상의 CDR, 가변 영역, 불변 영역, 불변 영역의 단편 또는 이들의 조합을 포함할 수 있다. 예시적인 LCF는 VL을 포함하고 경쇄 불변 영역("CL")의 전체 또는 일부를 포함한다. 문구 "LCF를 암호화하는 뉴클레오타이드 서열"은 LCF로 이루어지는 폴리펩티드를 암호화하는 뉴클레오타이드 서열 및 LCF를 함유하는 폴리펩티드(예를 들어, 특정 LCF 외에 추가 아미노산을 함유할 수 있는 폴리펩티드)를 암호화하는 뉴클레오타이드 서열을 포함한다. 예를 들어, LCF를 암호화하는 뉴클레오타이드 서열은, 특히, VL로 이루어지거나 전체 경쇄로 이루어지는 폴리펩티드를 암호화하는 뉴클레오타이드 서열을 포함한다.

[0076] 문구 "작동 가능하게 연결된"은 연결된 분자가 의도한대로 기능하는 방식으로 핵산 또는 단백질이 연결된 것을 지칭한다. DNA 영역은 이들이 서로 기능적으로 연관되어 있을 때 작동 가능하게 연결된다. 예를 들어, 프로모터는 상기 프로모터가 상기 서열의 전사에 참여할 수 있으면 암호화 서열에 작동가능하게 연결되고; 리보솜-결합 부위는 번역이 가능하도록 배치되는 경우 암호화 서열에 작동가능하게 연결된다. 일반적으로, 작동가능하게 연결된다는 것은 인접성을 포함하나, 반드시 필요하지는 않다. 분비 리더(leaders)와 같은 서열의 경우, 접근 및 해독 틀(reading frame)에서 적절한 배치는 일반적인 특징이다. 관심 유전자좌의 발현 강화 서열은 관심 유전자(GOI)와 기능적으로 연관되는 경우, 예를 들어, 발현 강화 서열이 존재함으로써 GOI의 발현이 강화되는 경우, GOI에 작동 가능하게 연결된다.

[0077] 관심 유전자좌, 예컨대 서열번호 1 또는 서열번호 2, 또는 이들의 단편을 기술할 때, "동일성 백분율"은 연속적인 상동성의 영역을 따라 인용된 동일성을 표시하는 상동 서열을 포함한다는 의미이지만, 비교된 서열에 상동성이 없는 갭, 결실 또는 삽입의 존재는 동일성 백분율의 계산에 고려되지 않는다.

[0078] 본원에서 사용된 바와 같이, 예컨대 종 동족체(homolog)를 갖는 서열번호 1 또는 이의 단편 사이의 "동일성 백분율" 결정은 종 동족체가 비교된 상동 서열을 정렬에서 갖지 않는 경우 서열 비교를 포함하지 않을 것이다(즉, 경우에 따라, 서열번호 1 또는 이의 단편은 그 지점에 삽입을 갖거나, 종 동족체는 갭 또는 결실을 갖는다). 따라서, "동일성 백분율"은 갭, 결실 및 삽입에 대한 불이익을 포함하지 않는다.

[0079] "인식 부위" 또는 "인식 서열"은 DNA 골격에 결합하고 DNA 골격의 부위 특이적인 절단을 유도하기 위해 뉴클레아제 또는 다른 효소에 의해 인식된 특정 DNA 서열이다. 엔도뉴클레아제는 DNA 분자 내에서 DNA를 절단한다. 인식 부위는 당업계에서 인식 표적 부위로도 지칭된다.

[0080] "재조합 효소 인식 부위(또는 "RRS")"는 재조합 효소, 예컨대 Cre 재조합 효소(Cre) 또는 플립파아제(flp)에 의해 인식되는 특정 DNA 서열이다. 부위 특이적인 재조합 효소의 하나 이상의 표적 인식 서열이 유기체의 게놈 내에 전략적으로 배치되는 경우, 부위 특이적인 재조합 효소는 결실, 도치 및 전좌를 비롯한 DNA 재배열을 수행할 수 있다. 하나의 예시에서, Cre는 8-bp 스페이서에 의해 분리된 2개의 13-bp 역위 반복으로 이루어진 이의 DNA 표적 인식 부위 *loxP*에서 재조합 사건을 특이적으로 매개한다. 예컨대, DNA의 재조합-매개 교환을 용이하게 하기 위해 하나 이상의 재조합 효소 인식 부위가 사용될 수 있다. 재조합 효소 인식 부위, 예컨대, *lox* 부위의 변이체 또는 돌연변이가 또한 사용될 수 있다(Araki, N. 등, 2002, *Nucleic Acids Research*, 30:19, e103).

[0081] "재조합 효소 매개 카세트 교환" 또는 "RMCE"는 게놈 표적 카세트를 공여자 카세트로 정확하게 치환하는 공정과 관련이 있다. 이러한 공정을 수행하기 위해 전형적으로 제공되는 분자 조성물은 (1) 특정 재조합 효소에 특이적인 인식 표적 부위에 의한 5' 및 3' 둘 다의 측부에 위치한 게놈 표적 카세트, (2) 인식 표적 부위의 매칭에 의해 측부에 위치한 공여자 카세트, 및 (3) 부위-특이적인 재조합 효소를 포함한다. 재조합 효소 단백질이 당업계에 널리 공지되어 있으며(Turan, S. 및 Bode J., 2011, *FASEB J.*, 25, pp. 4088-4107), 뉴클레오타이드의 수득 또는 손실 없이 특정 인식 표적 부위(DNA 서열) 내에서 DNA의 정확한 절단을 가능케 한다. 통상적인 재조합 효소/부위 조합은 비제한적으로 Cre/*lox* 및 Flp/*frrt*를 포함한다. 상업적으로 이용 가능한 키트도 R4-*attP* 부위를 함유하는 벡터 및 RMCE를 위해 phiC31 인테그라아제(integrase)를 암호화하는 벡터를 제공한다. (미국 공개 출원 제US20130004946호 또한 참조할 것.)

[0082] "부위 특이적 통합" 또는 "표적화된 삽입"은 유전자 또는 핵산 서열을 게놈상의 특정 위치에 직접 삽입 또는 통

합을 유도하도록, 즉, DNA를 인접한 폴리뉴클레오티드 사슬 내의 2개의 뉴클레오티드 사이의 특정 부위에 유도하도록 사용되는 유전자 표적화 방법을 지칭한다. 부위 특이적 통합 또는 표적화된 삽입은, 각각의 유전자가 자체 조절 요소(예컨대, 프로모터, 증강자 및/또는 전사적 말단 서열 등)를 갖는 다중 유전자와 같은 다중 발현 단위 또는 카세트를 포함하는 특정 핵산에 대해 수행될 수도 있다. "삽입" 및 "통합"은 상호교환적으로 사용된다. 유전자 또는 핵산 서열(예컨대, 발현 카세트를 포함하는 핵산 서열)의 삽입은 이용되는 유전자 편집 기법에 따라 하나 이상의 핵산의 치환, 또는 결실을 야기할 수 있다(또는 이를 위해 조작될 수 있다).

[0083] "안정적 통합"은 숙주 세포의 게놈 내에 통합된 외인성 핵산이 세포 배양물 내에서 연속된 기간 동안, 예를 들어, 적어도 7일, 적어도 10일, 적어도 15일, 적어도 20일, 적어도 25일, 적어도 30일, 적어도 35일, 적어도 40일, 적어도 45일, 적어도 50일, 적어도 55일 또는 적어도 60일 이상 통합상태를 유지하는 것을 의미한다. 대규모 제조 및 정제를 위해 이중 특이적 항원 결합 단백질을 만드는 것이 어려운 과제라는 것을 이해해야 한다. 안정성 및 복제 가능성은 임의의 생체분자의 재현성, 특히 치료학적으로 사용되는 생체분자의 재현성에 필수적이다. 본 개시의 방법에 의해 만들어진, 이중 특이적 항체를 발현하는 안정적 클론은 치료 생체분자를 생성하기 위한 일관되고 재현 가능한 방법을 제공한다.

[0084] 일반적인 설명

[0085] 본 개시는 숙주 세포 내의 다수의(예: 2개) 발현 강화 유전자좌를 사용하여 숙주 세포, 특히 차이나이즈 햄스터 (*Cricetulus griseus*) 세포주에서 다중 폴리펩티드의 발현을 개선하는 조성물 및 방법을 제공한다. 보다 구체적으로, 본 개시는, 항원 결합 단백질을 함께 암호화하는 다수의 외인성 핵산을 CHO 세포와 같은 숙주 세포에서 다수의 발현 강화 유전자좌 내에 부위 특이적 방식으로 통합하도록 설계된 조성물 및 방법을 제공한다. 특히, 본 개시는 다수의 발현 강화 유전자좌 내에 통합된 다수의 외인성 핵산을 함유하는 세포를 제공하되, 상기 다수의 외인성 핵산은 함께 항원 결합 단백질을 암호화한다. 본 개시는 다수의 외인성 핵산을 다수의 발현 강화 유전자좌 내에 부위 특이적으로 통합하도록 설계된 핵산 벡터를 추가로 제공한다. 본 개시는, 다수의 발현 강화 유전자좌 내에 벡터로부터의 다수의 외인성 핵산을 부위 특이적으로 통합하기 위해 다수의 발현 강화 유전자좌 각각에서 다수의 재조합 효소 인식 부위(RRS)를 함유하는 숙주 세포, 및 일치하는 RRS 및 다수의 외인성 핵산을 함유하는 벡터 세트를 포함하는 시스템을 추가적으로 제공한다. 또한, 본 개시는 본원에 개시된 세포, 벡터 및 시스템을 사용하여 항원 결합 단백질을 만드는 방법을 제공한다.

[0086] 다수의 발현 강화 유전자좌 내에 부위 특이적으로 통합된 다수의 외인성 핵산을 갖는 세포(Cells Having Multiple Exogenous Nucleic Acids Integrated Site-Specifically Within Multiple Expression Enhancing Loci)

[0087] 일 양태에서, 본 개시는 2개의 발현 강화 유전자좌에 부위 특이적으로 통합된 다수의 외인성 핵산을 함유하는 세포를 제공하며, 상기 다수의 외인성 핵산은 함께 항원 결합 단백질을 암호화한다. 항원 결합 단백질은 이중 특이적 항원 결합 단백질 또는 통상의(즉, 단일 특이적) 항원 결합 단백질일 수 있다.

[0088] 본원에 제공된 세포는 높은 역가로 및/또는 높은 특정 생산성(pg/세포/일)으로 원하는 항원 결합 단백질을 생산할 수 있다. 일부 구현예에서, 세포는 적어도 1 g/L, 1.5 g/L, 2.0 g/L, 2.5 g/L, 3.0 g/L, 3.5 g/L, 4.0 g/L, 4.5 g/L, 5.0 g/L, 또는 10 g/L 이상의 역가로 항원 결합 단백질을 생산한다. 일부 구현예에서, 항원 결합 단백질을 생산하는 세포는 일당 세포당 생산된 총 항원 결합 단백질(pg 단위)에 기초하여 결정된 적어도 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 피코그램/세포/일 이상의 특정 생산성을 갖는다.

[0089] 항원 결합 단백질을 함께 암호화하고 2개의 발현 강화 유전자좌 내에 통합된 외인성 핵산을 포함하는 숙주 세포는 생산 배양물 중에서 높은 세포 밀도(예: 1 내지 10×10^6 세포/mL)를 나타낸다. 다른 구현예에서, 항원 결합 단백질을 암호화하는 숙주 세포는 (생산 배양물 중에서) 적어도 5×10^6 세포/mL, 6×10^6 세포/mL, 7×10^6 세포/mL, 8×10^6 세포/mL, 9×10^6 세포/mL, 또는 10×10^6 세포/mL의 최종 세포 밀도에 도달한다.

[0090] 일부 구현예에서, 적어도 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 50% 또는 60% 이상의 이중 특이적 항원 결합 단백질 역가 대 총 항원 결합 단백질 역가의 비율로 이중 특이적 항원 결합 단백질을 생산할 수 있는 세포가 제공된다. 일부 구현예에서, 이중 특이적 항원 결합 단백질을 생산할 수 있는 세포로서, 이중 특이적 항원 결합 단백질 역가의 비율이 상기 세포에 의해 생산된 총 항원 결합 단백질 역가의 적어도 50%인 세포가 제공된다.

[0091] 다른 구현예에서, 항원 결합 단백질을 생산할 수 있는 세포로서, 하나의 유전자좌에서의 발현에 의해 생산된 총 항원 결합 단백질 역가와 비교하여 2개의 유전자좌에서의 발현에 의해 생산된 총 항원 결합 단백질 역가가 적어

도 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 50% 또는 60% 이상인 세포가 제공된다. 특정 구현예에서, 항원 결합 단백질을 생산할 수 있는 세포로서, 하나의 유전자좌에서의 발현에 의해 생산된 총 항원 결합 단백질 역가와 비교하여 2개의 유전자좌에서의 발현에 의해 생산된 총 항원 결합 단백질 역가가 적어도 0.5배, 0.75배, 1배, 1.5배, 1.75배 또는 2배 이상인 세포가 제공된다.

[0092] 일부 구현예에서, 세포는 제1 발현 강화 유전자좌 내의 특정 부위에서 통합된 제1 외인성 핵산, 및 제2 발현 강화 유전자좌 내의 특정 부위에서 통합된 제2 외인성 핵산을 함유하되, 제1 및 제2 외인성 핵산은 함께 항원 결합 단백질을 암호화한다. 제1 및 제2 외인성 핵산은 항원 결합 단백질의 HCF(들) 또는 LCF(들)(예: 가변 영역)을 암호화하는 다수의 뉴클레오타이드 서열을 함께 포함한다. 예를 들어, 단일 특이적 항체의 경우, 단일 특이적 항원 결합 단백질에 대해 HCF(예: VH)를 암호화하는 하나의 뉴클레오타이드 서열 및 LCF(예: VL)를 암호화하는 하나의 뉴클레오타이드 서열, 또는 각각의 다수의 카피(예: 2 카피)가 존재할 수 있다. 이중 특이적 항체의 경우, 각각 HCF를 암호화하는 2개의 뉴클레오타이드 서열(일반적으로 2개의 HCF는 서로 상이함), LCF를 암호화하는 뉴클레오타이드 서열의 1개 또는 2개의 카피, 또는 2개의 상이한 LCF를 암호화하는 2개의 뉴클레오타이드 서열이 존재할 수 있다. 항원 결합 단백질이 단일 특이적이거나 이중 특이적이거나 여부에 따라, HCF(들) 또는 LCF(들)(예: 가변 영역)을 암호화하는 뉴클레오타이드 서열은 2개의 발현 강화 유전자좌에서 상이한 변이체 또는 조합으로 통합될 수 있다. 예를 들어, 단일 특이적 항원 결합 단백질의 경우, 하나의 경우에, HCF를 암호화하는 뉴클레오타이드 서열 및 LCF를 암호화하는 뉴클레오타이드 서열은 2개의 유전자좌에 별도로, 각 유전자좌에 하나씩 통합될 수 있는 반면, 또 다른 경우에, HCF 및 LCF를 암호화하는 핵산이 하나의 유전자좌에 통합되고, 동일한 HCF 및 동일한 LCF를 암호화하는 별도의 핵산이 또 다른 유전자좌에 통합된다. 이중 특이적 항원 결합 단백질의 경우, 하나의 경우에, 제1 HCF와 LCF를 암호화하는 뉴클레오타이드 서열이 제1 유전자좌에 통합될 수 있고, 제2 HCF를 암호화하는 뉴클레오타이드 서열이 제2 유전자좌에 통합되되, 2개의 HCF는 상이하고, LCF는 이중 특이적 항원 결합 단백질의 공통 LCF이며; 또 다른 경우에, 제1 HCF와 LCF를 암호화하는 뉴클레오타이드 서열이 제1 유전자좌에 통합되고, 제2 HCF 및 동일한 LCF를 암호화하는 뉴클레오타이드 서열이 제2 유전자좌에 통합된다.

[0093] 일부 구현예에서, HCF 또는 LCF를 암호화하는 뉴클레오타이드 서열은 불변 영역으로부터 아미노산을 암호화할 수 있다. 예를 들어, HCF 또는 LCF를 암호화하는 뉴클레오타이드 서열은 CL, CH1, CH2, CH3, 또는 CH1, CH2 또는 CH3의 조합 중 하나 이상을 암호화할 수 있거나, 전체 불변 영역을 암호화할 수 있다. 일부 구현예에서, HCF를 암호화하는 뉴클레오타이드 서열은 CH3 도메인을 암호화한다. 특정 구현예에서, HHCF를 암호화하는 뉴클레오타이드 서열은 중쇄를 암호화한다. 일부 구현예에서, LCF를 암호화하는 뉴클레오타이드 서열은 경쇄를 암호화한다.

[0094] 2개의 HCF가 포함된 구현예에서, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 불변 영역으로부터 아미노산을 암호화할 수 있고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 불변 영역으로부터 아미노산을 암호화할 수 있으며, 2개의 불변 영역으로부터의 아미노산은 동일하거나 적어도 하나의 위치(예컨대, 상이한 단백질 A 결합 특성을 나타내는 위치들, 또는 다양한 이중 특이적 항원 결합 단백질에 대해 아래 본원에 기술된 다른 위치들)가 상이할 수 있다. 아미노산이 상이한 것과는 별개로, 2개의 불변 영역으로부터 아미노산을 암호화하는 2개의 뉴클레오타이드 서열은 하나의 뉴클레오타이드 서열의 하나 이상의 코돈을 변형함으로써 구별될 수 있는데, 이는 핵산에 기초한 검정에서 2개의 뉴클레오타이드 서열을 구별하는 편리한 기초를 제공한다.

[0095] 일부 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 프로모터를 함유하는 전사 조절 서열에 독립적으로 및 작동 가능하게 연결된다. "독립적으로"는, 각각의 코딩 서열이 프로모터와 같은 별도의 전사 조절 서열에 작동 가능하게 연결되어, 코딩 서열의 전사가 별도로 조절 및 제어된다는 것을 의미한다. 일부 구현예에서, 2개의 HCF 함유 폴리펩티드의 전사를 유도하는 프로모터들은 동일하다. 일부 구현예에서, 2개의 HCF 함유 폴리펩티드의 전사를 유도하는 프로모터들뿐만 아니라 VL 함유 폴리펩티드의 전사를 유도하는 프로모터(예: CMV 프로모터)도 모두 동일하다. 일부 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 유도 프로모터 또는 억제 프로모터에 독립적으로 및 작동 가능하게 연결된다. 유도 프로모터 및 억제 프로모터는 생산 단계(유가 배양)에서만 생산이 발생하도록 하고 성장 단계(시드 트레인 배양) 동안에는 발생하지 않도록 하거나, 상이한 유전자좌에서 항체 성분(HCF 및 LCF)의 발현을 정밀하게 상이하게 조절할 수 있게 한다. 각각의 유전자 산물의 생산(발현)의 미세 조절은 상이한 프로모터에 의해 달성될 수 있다.

[0096] 하나의 이러한 실시예에서, 세포가 테트라시클린 억제 단백질(TetR)을 발현하도록 먼저 조작되고, 각각의 HCF 및 LCF 코딩 뉴클레오타이드 서열을 TetR에 의해 활성이 조절되는 프로모터의 전사 조절 하에 놓인다. 2개의 탠덤 TetR 작동자(TetO)는 CMV 프로모터의 바로 하류에 위치한다. 일부 구현예에서, 각각의 HCF 및/또는 LCF 코딩 뉴클레오타이드 서열은 적어도 하나의 TetR 작동자(TetO) 또는 Arc 작동자(ArcO)의 상류에서 프로모터에 독립적으로 및 작동 가능하게 연결된다. 일부 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 CMV/TetO 또는

CMV/ArcO 하이브리드 프로모터에 독립적으로 및 작동 가능하게 연결된다. 적절한 추가 프로모터는 아래 본원에 기술된다.

- [0097] 일부 구현예에서, 2개의 유전자좌에 통합된 다수의 외인성 핵산은 RRS가 측면에 위치한다. 예를 들어, 제1 RRS 및 제2 RRS는 제1 유전자좌에 통합된 제1 외인성 핵산에 대해 상대적으로 5' 및 3'에 각각 위치되고, 제3 RRS 및 제4 RRS는 제2 유전자좌에 통합된 제2 외인성 핵산에 대해 상대적으로 5' 및 3'에 각각 위치되며, 제1 및 제2 RRS가 상이하고, 제3 및 제4 RRS가 상이하다. 일부 구현예에서, 제1, 제2, 제3 및 제4 RRS는 모두 서로 상이하다. 다른 구현예에서, 제1 외인성 핵산이 HCF 및 LCF를 암호화하고, 제2 외인성 핵산이 동일한 HCF 및 LCF를 암호화하는 경우, 제1 및 제3 RRS는 동일하고, 제2 및 제4 RRS는 동일하다.
- [0098] 일부 구현예에서, 일 유전자좌에 통합된 외인성 핵산이 2개의 HCF 또는 LCF 코딩 뉴클레오티드 서열을 포함하는 경우, 추가 RRS가 2개의 뉴클레오티드 사이에 포함될 수 있다. 이러한 추가 RRS는 외인성 핵산의 측면에 위치하는 2개의 RRS와는 상이한 것이어야 한다. 일부 구현예에서, 추가 RRS는, 통합된 외인성 핵산에 포함되어 2개의 HCF 또는 LCF 코딩 뉴클레오티드 서열 사이에 위치한 선별 표지 유전자의 인트론 내에 삽입된다. 전사 및 전사 후 변형(post transcriptional processing) 이후에, 인트론이 절단되어 선별 표지를 암호화하는 mRNA가 생성될 것이다. 제1 유전자좌에 통합된 외인성 핵산 및 제2 유전자좌에 통합된 외인성 핵산 모두가 2개의 HCF 또는 LCF 코딩 서열(예: LCF-HCF1 및 LCF-HCF2)을 포함하는 구현예에서, 각각의 유전자좌에서 2개의 HCF 또는 LCF 코딩 서열 사이의 제1 및 제2 외인성 핵산 중 하나에만 또는 둘 모두에 추가 RRS가 포함될 수 있다. 제1 유전자좌에서의 추가 RRS는 제2 유전자좌에서의 추가 RRS와 동일하거나 상이할 수 있다. 각각의 추가 RRS는 해당 유전자좌에 통합된 외인성 핵산의 측면에 위치하는 2개의 RRS와는 상이한 것이어야 한다. 각각의 추가 RRS는 선택적으로 선별 표지 유전자의 인트론 내에 삽입될 수 있고, 상기 2개의 유전자좌 내에 인트론을 갖는 선별 표지 유전자들은 상이할 수 있다.
- [0099] 다수의 HCF 또는 LCF 코딩 서열이 일 유전자좌에 통합된 외인성 핵산 내에 포함되는 경우, 유전자좌 내의 다수의 코딩 서열의 상대 위치는 달라질 수 있다. 예를 들어, 통합된 외인성 핵산이 LCF 암호화 뉴클레오티드 서열 및 HCF 암호화 뉴클레오티드 서열을 포함하는 구현예에서, LCF 암호화 뉴클레오티드 서열은 HCF 암호화 뉴클레오티드 서열에 대해 상대적으로 상류 또는 하류에 위치할 수 있다. 특정 구현예에서, LCF 암호화 뉴클레오티드 서열은 HCF 암호화 뉴클레오티드 서열에 대해 상대적으로 상류에 위치한다. 두 유전자좌 모두가 LCF 암호화 뉴클레오티드 서열 및 HCF 암호화 뉴클레오티드 서열을 포함하는 경우, 특정 구현예에서, 두 유전자좌 모두에서 LCF 암호화 뉴클레오티드 서열은 HCF 암호화 뉴클레오티드 서열에 대해 상대적으로 상류에 위치한다.
- [0100] 다른 구현예에서, 제1 발현 강화 유전자좌 내에 통합된 제1 RRS 쌍 및 제2 발현 강화 유전자좌 내에 통합된 제2 RRS 쌍을 포함하는 세포로서, 각 쌍 내의 2개의 RRS는 상이한 세포가 제공된다. 이러한 세포는, 항원 결합 단백질을 함께 암호화하는 다수의 통합 대상 외인성 핵산을 수용하는 데 유용하다.
- [0101] 일부 구현예에서, 제1 외인성 핵산은 제1 유전자좌에서 2개의 RRS 사이에 존재하고, 제2 외인성 핵산은 제2 유전자좌에서 2개의 RRS 사이에 존재한다. 제1 및 제2 외인성 핵산은 각각 하나 이상의 선별 표지 유전자를 암호화할 수 있다. 선별 표지 유전자는 서로 상이할 수 있다.
- [0102] 일부 구현예에서, 추가 RRS가 유전자좌에서 쌍을 이루는 2개의 RRS (즉, 5' RRS 및 3' RRS) 사이에 존재하되, 추가 RRS는 유전자좌에서의 5' RRS 및 3' RRS 둘 모두와 상이하다. 일부 구현예에서, 추가 RRS가 2개의 유전자좌 중 하나에서의 5' RRS와 3' RRS 사이에 존재하고, 다른 구현예에서, 추가 RRS는 2개의 유전자좌 각각에서의 5' RRS와 3' RRS 사이에 존재한다. 추가 RRS가 5' RRS와 3' RRS 사이에 존재하는 경우, 선별 표지 유전자는 5' RRS와 추가 RRS 사이에 포함될 수 있고, 또 다른 상이한 선별 표지 유전자가 추가 RRS와 3' RRS 사이에 포함될 수 있으며, 2개의 선별 표지는 상이하다.
- [0103] 기술된 많은 구현예에서, 세포는 CHO 세포이며, 2개의 강화된 발현 유전자좌 중 1개는 서열번호 1과 적어도 90% 동일한 뉴클레오티드 서열, 서열번호 2와 적어도 90% 동일한 뉴클레오티드 서열, 및 서열번호 3과 적어도 90% 동일한 뉴클레오티드 서열로 이루어지는 군으로부터 선택된다.
- [0104] **이중 특이적 항원 결합 단백질(Bispecific Antigen-Binding Proteins)**
- [0105] 본 개시에 기술된 세포, 벡터 및 시스템에서의 클로닝 및 생산에 적합한, 이중 특이적 항체와 같은 이중 특이적 항원 결합 단백질은 이중 특이적 항원 결합 단백질의 임의의 특정 포맷으로 한정되지 않는다.
- [0106] 다양한 구현예에서, 이중 특이적 항원 결합 단백질은 2개의 폴리펩티드(각각 HC와 같은 항원 결합 모이어티를 함유함) 및 CH3 도메인을 포함하며, 2개의 폴리펩티드의 항원 결합 모이어티는 상이한 항원 특이성을 가지며, 2

개의 CH3 도메인은 CH3 도메인 중 하나가 적어도 하나의 아미노산 위치가 변형되어 2개의 폴리펩티드 사이에서 분별 단백질 A 결합 특성을 생성한다는 점에서 서로에 대해 다른 이종 이량체이다. 예를 들어, 미국 특허 제 8,586,713호에 기술된 이종 특이적 항체를 참조할 것. 이러한 방식으로, 분별 단백질 A 단리법을 사용하여 동종 이량체로부터 이종 이량체 이종 특이적 항원 결합 단백질을 쉽게 단리할 수 있다.

[0107] 일부 구현예에서, 이종 특이적 항원 결합 단백질은, 상이한 항원 특이성을 가지고 CH3 도메인에서 적어도 하나의 아미노산 위치가 상이한 2개의 중쇄를 포함하여, 2개의 중쇄 사이의 분별 단백질 A 결합 특성을 생성한다.

[0108] 일부 구현예에서, 2개의 폴리펩티드는 인간 IgG의 CH3 도메인을 함유하되, 2개의 폴리펩티드 중 하나는 IgG1, IgG2 및 IgG4로부터 선택된 인간 IgG의 CH3 도메인을 함유하고, 2개의 폴리펩티드 중 다른 하나는 IgG1, IgG2 및 IgG4로부터 선택된 인간 IgG의 변형된 CH3를 함유하며, 상기 변형은 단백질 A에 대한 변형된 CH3 영역의 결합을 감소시키거나 제거한다. 특정 구현예에서, 2개의 폴리펩티드 중 하나는 인간 IgG1의 CH3 도메인을 함유하고, 2개의 폴리펩티드 중 다른 하나는 인간 IgG1의 변형된 CH3 도메인을 함유하되, 상기 변형은 IMGT 엑손 넘버링 시스템에서의 (i) 95R 및 (ii) 95R과 96F로 이루어지는 군으로부터 선택된다. 다른 특정 구현예에서, 변형된 CH3 도메인은 IMGT 엑손 넘버링 시스템에서의 16E, 18M, 44S, 52N, 57M, 및 82I로 이루어지는 군으로부터 선택된 1개 내지 5개의 추가 변형을 포함한다.

[0109] 다른 다양한 구현예에서, 2개의 폴리펩티드는 마우스 IgG의 CH3 도메인을 함유하되, 2개의 폴리펩티드 중 하나는 변형되지 않은 마우스 IgG의 CH3 도메인을 함유하고, 2개의 폴리펩티드 중 다른 하나는 마우스 IgG의 변형된 CH3 도메인을 함유하며, 상기 변형은 단백질 A에 대한 변형된 CH3 영역의 결합을 감소시키거나 제거한다. 다양한 구현예에서, 마우스 IgG CH3 영역이 변형되어, 다음으로 이루어지는 군으로부터 선택된 특정 위치(EU 넘버링)에 특정 아미노산을 포함한다: 252T, 254T, 및 256T; 252T, 254T, 256T, 및 258K; 247P, 252T, 254T, 256T, 및 258K; 435R 및 436F; 252T, 254T, 256T, 435R, 및 436F; 252T, 254T, 256T, 258K, 435R, 및 436F; 247P, 252T, 254T, 256T, 258K, 435R, 및 436F; 및 435R. 특정 구현예에서, 다음으로 이루어지는 군으로부터 선택된 특정 변형 군이 만들어진: M252T, S254T, S256T; M252T, S254T, S256T, I258K; I247P, M252T, S254T, S256T, I258K; H435R, H436F; M252T, S254T, S256T, H435R, H436F; M252T, S254T, S256T, I258K, H435R, H436F; I247P, M252T, S254T, S256T, I258K, H435R, H436F; 및 H435R.

[0110] 다양한 구현예에서, 이종 특이적 항원 결합 단백질은 마우스 및 랫트 단클론 항체 또는 항원 결합 단백질의 하이브리드, 예를 들어 마우스 IgG2 및 랫트 IgG2b의 하이브리드이다. 이들 구현예에 따르면, 이종 특이적 항체는 하나의 중쇄 쌍/경쇄 쌍 각각을 포함하는 2개의 항체 이종 이량체로 이루어지며, 이들의 Fc 부분을 통해 결합한다. 원하는 이종 이량체는 2개의 부모 항체 동종 이량체 및 이종 특이적 이종 이량체의 혼합물로부터 용이하게 정제될 수 있는데, 이는 단백질 A에 대한 이종 특이적 항체의 결합 속성이 부모 항체의 결합 속성과 다르고, 따라서 랫트 IgG2b는 단백질 A에 결합하지 않는 반면 마우스 IgG2a는 단백질 A에 결합하기 때문이다. 결과적으로, 마우스-랫트 이종 이량체는 단백질 A에 결합하지만, 마우스 IgG2a 동종 이량체보다 더 높은 pH에서는 용출되는데, 이는 이종 특이적 이종 이량체의 선택적 정제를 가능하게 한다.

[0111] 다른 다양한 구현예에서, 이종 특이적 항원 결합 단백질은 당업계에서 "노브-인투-홀(knobs-into-holes)"로 지칭되는 포맷이다(예: 미국 특허 제7,183,076호 참조). 이들 구현예에서, 2개의 항체의 Fc 부분은 조작되어, 이 중 하나는 돌출 "노브"가 되고 다른 하나는 상보성 "홀"을 제공한다. 동일한 세포에서 생산될 때, 중쇄는 조작된 "노브"를 조작된 "홀"과 결합시킴으로써 동종 이량체 보다는 이종 이량체를 우선적으로 형성하는 것으로 알려져 있다.

[0112] 또 다른 구현예에서, 제1 중쇄 및 제2 중쇄는 2개의 중쇄 사이의 상호 작용이 가능하도록 CH3 도메인에서 하나 이상의 아미노산 변형을 포함한다. CH3-CH3 계면 아미노산 잔기는 하전된 아미노산과 대체되어 정전기적으로 부적합한 동종 이량체를 형성할 수 있다. (예: PCT 공보 제W02009089004호, 및 유럽 공보 제EP1870459호 참조)

[0113] 다른 구현예에서, 이종 이량체의 우선 형성을 촉진하기 위해 제1 중쇄는 동형 IgA의 CH3 도메인을 포함하고, 제2 중쇄는 IgG의 CH3 도메인을 포함한다(또는 그 반대이다). (예: PCT 공보 제W02007110205호 참조)

[0114] 다른 구현예에서, Fab-아암 교환(PCT 공보 제W02008119353호; PCT 공보 제W02011131746호), 꼬인 나선 도메인 상호 작용(PCT 공보 제W02011034605호) 또는 류신 지퍼 펩티드(leucine zipper peptides)(Kostelny, 외의 *J. Immunol.* 1992, 148(5):1547-1553)와 같은 이종 이량체의 형성을 촉진시키는 조작 방법에 의해 다양한 포맷이 면역 글로불린 사슬과 통합될 수 있다.

[0115] 이종 특이적 항원 결합 단백질을 생성하는 데 사용될 수 있는 면역 글로불린 중쇄 단편(예: 가변 영역)은 당업

계에 공지된 임의의 방법을 사용해 생성할 수 있다. 예를 들어, 제1 중쇄는 제1 항원으로 면역화된 제1 동물의 성숙한 B 세포의 계놈으로부터 유래된 핵산에 의해 암호화되는 가변 영역을 포함하고, 제1 중쇄는 제1 항원을 특이적으로 인식하며; 제2 중쇄는 제2 항원으로 면역화된 제2 동물의 성숙한 B 세포로부터 유래된 핵산에 의해 암호화되는 가변 영역을 포함하고, 제2 중쇄는 제2 항원을 특이적으로 인식한다. 면역 글로불린 중쇄 가변 영역 서열은 파지 디스플레이와 같은, 당업계에서 공지된 임의의 다른 방법에 의해서도 획득될 수 있다. 다른 실시예에서, 중쇄 가변 영역을 암호화하는 핵산에는 당업계에서 기술되었거나 달리 이용 가능한 항체의 핵산이 포함된다. 일부 구현예에서, 2개의 중쇄 코딩 서열 중 하나는 핵산 기반의 검정에서 2개의 코딩 서열을 구별하기 위한 편리한 기초를 제공하기 위해 코돈 변형된 것이다.

[0116] 2개의 상이한 에피토프(또는 2개의 상이한 항원)를 인식하는 2개의 중쇄를 포함하는 이중 특이적 항체는 이들이 동일한 경쇄(즉, 동일한 가변 및 불변 도메인을 갖는 경쇄)와 쌍을 이룰 수 있는 곳에서 더 쉽게 분리된다. 예를 들어, 미국 특허 제8,586,713호 및 본원에 기술된 기술분야에서 설명된 바와 같이, 중쇄 가변 도메인과 그의 표적 항원 간의 선택도 및/또는 친화성을 방해하지 않거나 실질적으로 방해하지 않으면서 특이성이 상이한 2개의 중쇄와 쌍을 이룰 수 있는 경쇄를 생성하는 다양한 방법이 당업계에 공지되어 있다.

[0117] 이중 특이적 항원 결합 단백질은 이중 항원 특이성 및 관련된 유용한 용도를 가질 수 있다.

[0118] 일부 실시예에서, 종양 항원 및 T 세포 항원에 대한 결합 특이성을 갖는 이중 특이적 항원 결합 단백질은 세포 상의 항원(예: CD20)을 표적화하고, T 세포 상의 항원(예: CD3과 같은 T 세포 수용체)를 또한 표적화하도록 만 들어질 수 있다. 이러한 방식으로, 이중 특이적 항원 결합 단백질은 환자의 T 세포뿐만 아니라 환자에서의 관심 세포(예: CD20 결합을 통한 림프종 환자에서의 B 세포)를 표적화한다. 다양한 구현예에서, 이중 특이적 항원 결합 단백질은 CD3에 결합할 때 T 세포를 활성화시켜 T 세포 활성화를 특정한 선택된 종양 세포에 대한 결합시키도록 설계된다.

[0119] 하나의 모이어티가 T 세포 수용체에 결합하고(예컨대 CD3에 결합하고) 다른 하나가 표적 항원에 결합하는 이중 특이적 항원 결합 단백질의 맥락에서, 표적 항원은 종양 관련 항원일 수 있다. 특이적인 종양 관련 항원의 비한정적인 예에는, 예를 들어, AFP, ALK, BAGE 단백질, BIRC5 (서바이빈), BIRC7, β -카테닌, bcr-abl, BRCA1, BCMA, BORIS, CA9, 탄산무수화효소 IX, 카스파제-8, CALR, CCR5, CD19, CD20(MS4A1), CD22, CD30, CD40, CDK4, CEA, CLEC-12, CTLA4, 사이클린-B1, CYP1B1, EGFR, EGFRvIII, ErbB2/Her2, ErbB3, ErbB4, ETV6-AML, EpCAM, EphA2, Fra-1, FOLR1, GAGE 단백질(예: GAGE-1, -2), GD2, GD3, GloboH, 글리피칸-3, GM3, gp100, Her2, HLA/B-raf, HLA/k-ras, HLA/MAGE-A3, hTERT, LMP2, MAGE 단백질(예: MAGE-1, -2, -3, -4, -6, 및 -12), MART-1, 메소텔린, ML-IAP, Muc1, Muc2, Muc3, Muc4, Muc5, Muc16 (CA-125), MUM1, NA17, NY-BR1, NY-BR62, NY-BR85, NY-ESO1, OX40, p15, p53, PAP, PAX3, PAX5, PCTA-1, PLAC1, PRLR, PRAME, PSMA (FOLH1), RAGE 단백질, Ras, RGS5, Rho, SART-1, SART-3, Steap-1, Steap-2, TAG-72, TGF- β , TMPRSS2, 톨슨-누벨 항원(Tn), TRP-1, TRP-2, 티로시나제, 및 유로플라킨-3이 포함된다. 일부 구현예에서, 이중 특이적 항원 결합 단백질은 CD3에 결합하는 하나의 모이어티를 포함한다. 예시적인 항CD3 항체 모이어티는 미국 특허 출원 공개 번호 제 US2014/0088295A1호 및 제US20150266966A1호, 및 2017년 3월 30일 공개되고 그 전체가 본원에 참조로서 통합된 국제 공보 제WO 2017/053856호에 기술되어 있다. 다른 구현예에서, 이중 특이적 항원 결합 단백질은 CD3에 결합하는 하나의 모이어티 및 BCMA, CD19, CD20, CD28, CLEC-12, Her2, HLA 단백질, MAGE 단백질, Muc16, PSMA, 또는 Steap-2에 결합하는 하나의 모이어티를 포함한다. 또 다른 구현예에서, 이중 특이적 항원 결합 단백질은 (본원에 참조로서 통합된 미국 특허 출원 공개 번호 제US2014/0088295A1호 및 제US20150266966A1에 기술된 바와 같은) 항CD3 x 항CD20 이중 특이적 항체, 항CD3 x 항 Mucin 16 이중 특이적 항체(예: 항CD3 x 항Muc16 이중 특이적 항체), 및 항CD3 x 항전립선 특이적 막 항원 이중 특이적 항체(예: 항CD3 x 항-PSMA 이중 특이적 항체)를 포함한다.

[0120] 하나의 모이어티가 T 세포 수용체에 결합하고(예컨대 CD3에 결합하고) 다른 하나가 표적 항원에 결합하는 이중 특이적 항원 결합 단백질의 맥락에서, 표적 항원은 감염성 질병 관련 항원일 수 있다. 감염성 질환 관련 항원의 비한정적인 예에는, 예를 들어, 바이러스 입자의 표면에서 발견되거나 바이러스에 감염된 세포 상에서 우선적으로 발견되는 항원이 포함되며, 상기 바이러스는 HIV, 간염 (A, B 또는 C), 헤르페스 바이러스 (예: HSV-1, HSV-2, CMV, HAV-6, VZV, 엡스타인 바 바이러스), 아데노 바이러스, 인플루엔자 바이러스, 플라비 바이러스, 에코 바이러스, 코감기 바이러스, 콕사키 바이러스(coxsackie virus), 코로나 바이러스, 호흡기 세포융합 바이러스, 이차선염 바이러스, 로타 바이러스, 홍역 바이러스, 풍진 바이러스, 파보 바이러스, 우두 바이러스, HTLV, 뎅기 바이러스, 유두종 바이러스, 연체 동물 바이러스, 폴리오 바이러스, 광견병 바이러스, JC 바이러스 및 아르보 바이러스성 뇌염 바이러스로 이루어지는 군으로부터 선택된다. 대안적으로, 표적 항원은 박테리아의 표면에서

발현되거나, 박테리아에 감염된 세포 상에서 우선적으로 발현되는 항원일 수 있으며, 상기 박테리아는 클라미디아(chlamydia), 리케차(rickettsia), 마이코 박테리아(mycobacteria), 포도상 구균, 연쇄상 구균, 폐렴 구균, 수막 구균, 임균, 클렙시엘라(klebsiella), 프로테우스(proteus), 세라티아(serratia), 슈도모나스(pseudomonas), 레지오넬라(legionella), 디프테리아(diphtheria), 살모넬라균(salmonella), 간균(bacilli), 콜레라(cholera), 과상풍(tetanus), 보툴리누스 중독(botulism), 탄저병(anthrax), 역병(plague), 렙토스피라(leptospira), 및 라임 병원균(Lyme disease bacteria)로 이루어지는 군으로부터 선택된다. 특정 구현예에서, 표적 항원은 곰팡이의 표면에서 발현되거나 곰팡이에 감염된 세포 상에서 우선적으로 발현되는 항원이며, 상기 곰팡이는 칸디다(알비칸스, 크루세이, 글라브라타, 트리파칼리스 등), 크립토코커스 네오포르만스(Cryptococcus neoformans), 아스페리길루스(푸미가투스, 니제르 등), 털곰팡이목(털곰팡이속, 활털곰팡이속, 거미줄 곰팡이속 등), 스포로트릭스 쉐키아이(Sporothrix schenckii), 블라스토미세스 더마티티디스(Blastomyces dermatitidis), 파라코키시오이디즈 브라질리엔시스(Paracoccidioides brasiliensis), 콕시디오이데스 이미티스(Coccidioides immitis), 및 히스토플라스마속(Histoplasma capsulatum)으로 이루어지는 군으로부터 선택된다. 특정 구현예에서, 표적 항체는 기생충의 표면 상에서 발현되거나 기생충에 감염된 세포 상에서 우선적으로 발현되는 항원이며, 상기 기생충은 이질 아메바(Entamoeba histolytica), 대장 발란티듐(Balantidium coli), 파울러 자유아메바(Naegleria fowleri), 가시 아메바종(Acanthamoeba sp.), 람블 편모충(Giardia lamblia), 크립토스포리디움종(Cryptosporidium sp.), 주폐포자충(Pneumocystis carinii), 삼일열원충(Plasmodium vivax), 바베시아 미코티(Babesia microti), 트리파노소마 브루세이(Trypanosoma brucei), 트리파노소마 크루지(Trypanosoma cruzi), 리슈만 편모충(Leishmania donovani), 톡소포자충(Toxoplasma gondii), 쥐모양 선충(Nippostrongylus brasiliensis), 붉은 여우 조충(Taenia crassiceps), 및 말레이 사상충(Brugia malayi)으로 이루어지는 군으로부터 선택된다. 특정한 병원균 관련 항원의 비한정적 예에는, 예를 들어, HIV gp120, HIV CD4, B형 간염 당단백질 L, B형 간염 당단백질 M, B형 간염 당단백질 S, C형 간염 E1, C형 간염 E2, 간세포 특이적 단백질, 단순 포진 바이러스 gB(herpes simplex virus gB), 시토크로미 바이러스 gB(cytomegalovirus gB), 및 HTLV 외피 단백질(HTLV envelope protein)이 포함된다.

[0121] 동일한 세포의 표면 상에서 각각이 결합 파트너를 향하는 (즉, 각각이 상이한 표적을 향하는) 2개의 결합 모이어티를 포함하는 이중 특이적 결합 단백질도 만들 수 있다. 이러한 디자인은 동일한 세포의 표면 상에서 표적 모두를 발현하는 특정 세포 또는 세포 유형을 표적으로 하기에 특히 적합하다. 표적이 다른 세포 상에서 개별적으로 나타날 수 있다 하더라도, 이들 결합 단백질의 결합 모이어티는 각각의 결합 모이어티가 비교적 낮은 (예를 들어, 낮은 나노몰 또는 가령, 500, 600, 700, 800 나노몰과 같은 100 나노몰 KD보다 큰 높은 나노몰) 친화성으로 그의 표적에 결합하도록 선택된다. 이러한 방식으로, 장기간의 표적 결합은 2개의 표적이 동일한 세포 상에서 근접한 상황에서만 선호된다.

[0122] 각각 동일한 표적의 상이한 에피토프에서 동일한 표적에 결합하는 2개의 결합 모이어티를 포함하는 이중 특이적 결합 단백질을 만들 수 있다. 이러한 디자인은 결합 단백질로 표적을 성공적으로 차단하는 확률을 최대화하는데 특히 적합하다. 예를 들어, 막관통 경로 또는 세포 표면 수용체의 다수의 세포 외 루프는 동일한 이중 특이적 결합 분자에 의해 표적화될 수 있다.

[0123] 면역 신호 전달의 음성 조절체를 모으고 활성화시켜 면역을 억제시키는 2개의 결합 모이어티를 포함하는 이중 특이적 결합 단백질을 만들 수 있다. 표적이 동일한 세포 상에 있는 경우 cis의 억제가 달성될 수 있고, 표적이 상이한 세포 상에 있는 경우 trans의 억제가 달성될 수 있다. 예를 들어, cis의 억제는 항IgGRIIb 결합 모이어티 및 항FeID1 결합 모이어티를 갖는 이중 특이적 결합 단백질로 달성될 수 있으므로, IgGRIIb는 FeID1에 대한 면역 반응을 하향 조절하기 위해 FeID1의 존재 하에서만 모인다. 예를 들어, trans의 억제는 항BTLA 결합 모이어티 및 조직 특이적 관심 항원에 특이적으로 결합하는 결합 모이어티를 갖는 이중 특이적 결합 단백질로 달성될 수 있으므로, 억제성 BTLA 분자의 클러스터링은 자기 면역 질병을 잠재적으로 다루는 선택된 표적 조직에서만 발생한다.

[0124] 다중 성분 수용체를 활성화시키는 이중 특이적 결합 단백질을 만들 수 있다. 이러한 디자인에서, 수용체의 2개의 성분을 향하는 2개의 결합 모이어티는 수용체에 결합하고, 가교 결합하며, 수용체로부터의 신호 전달을 활성화한다. (결합이 수용체를 가교 결합시키는 경우) 이는, 예를 들어, IFNAR1에 결합하는 결합 모이어티 및 IFNAR2에 결합하는 결합 모이어티를 갖는 이중 특이적 결합 단백질을 사용해 수행될 수 있다. 이러한 이중 특이적 결합 단백질은 인터페론 치료에 대한 대안을 제공할 수 있다.

[0125] 반투과성 배리어, 예를 들어, 혈액-뇌 배리어를 가로질러 결합 모이어티를 운반하는 이중 특이적 결합 단백질을 만들 수 있다. 이러한 디자인에서, 하나의 결합 모이어티는 특정 선택성 배리어를 통과할 수 있는 표적에 결합

하고, 다른 하나의 모이어티는 치료 활성을 갖는 분자를 표적화하되, 치료 활성을 갖는 표적 분자는 배리어를 정상적으로 통과할 수 없다. 이러한 종류의 이중 특이적 결합 단백질은 치료제가 달리 도달할 수 없는 조직에 치료제를 전달하는데 유용하다. 일부 실시예는 장 또는 폐 내에 치료제를 운반하기 위해 pIGR 수용체를 표적화하거나, 혈액-뇌 배리어를 가로질러 치료제를 운반하기 위해 트랜스페린 수용체를 표적화하는 것을 포함한다.

[0126] 특이적 세포 또는 세포 유형 내에 결합 모이어티를 운반하는 이중 특이적 결합 단백질을 만들 수 있다. 이러한 디자인에서, 하나의 결합 모이어티는 세포 내로 쉽게 내부화되는 세포 표면 단백질(예: 수용체)을 표적화한다. 다른 하나의 모이어티는 치료 효과를 일으키는 세포간 단백질을 표적화한다.

[0127] 대식 면역 세포의 표면 수용체 및 감염성 병원균의 표면 분자(예: 효모 또는 박테리아)에 결합하여 대식 면역 세포의 근위에 감염성 병원균을 운반함으로써 병원균의 식균 작용을 용이하게 하는 이중 특이적 결합 단백질. 이러한 디자인의 예는 CD64 또는 CD89 분자 및 병원균을 또한 표적화하는 이중 특이적 항체일 것이다.

[0128] 항체 가변 영역을 하나의 결합 모이어티로서 가지고 비-Ig 모이어티를 제2 결합 모이어티로서 갖는 이중 특이적 결합 단백질. 항체 가변 영역은 표적화를 달성하는 반면, 비-Ig 모이어티는 Fc에 연결된 작동자(effector) 또는 독소(toxin)이다. 이러한 방식으로, 리간드(예: 작동자 또는 독소)가 항체 가변 영역에 의해 결합된 표적에 전달된다.

[0129] 임의의 2개의 단백질 모이어티가 Fc의 맥락에서 서로의 근위에 전달되도록 Ig 영역(예: CH2 및 CH3 영역을 함유하는 Ig 서열)에 각각 결합된 2개의 모이어티를 가지는 이중 특이적 결합 단백질. 이러한 디자인의 예는 트랩, 예를 들어 동종 또는 이종 이량체 트랩 분자가 포함된다.

[0130] 발현 강화 유전자좌(Expression-Enhancing Loci)

[0131] 본 발명에서의 용도에 적합한 발현 강화 유전자좌는, 예를 들어, 서열번호 1에 대해 실질적 상동성을 갖는 뉴클레오티드 서열을 갖는, 미국 특허 제8,389,239호에 기술된 바와 같은, 유전자좌("EESYR[®] 유전자좌" 또는 "1번 유전자좌"로서도 본원에 지칭됨), 서열번호 2 또는 서열번호 3에 대해 실질적 상동성을 갖는 뉴클레오티드 서열을 포함하는, 미국 출원 일련 번호 제14/919,300호에 기술된 바와 같은, 유전자좌("YARS 유전자좌" 또는 "2번 유전자좌"로서도 본원에 지칭됨), 및 당업계에서 문서화된 (예: US 20150167020A1, 및 미국 특허 제6,800,457호) 다른 발현 강화 유전자좌 및 서열을 포함한다.

[0132] 일부 구현예에서, 본 발명에 사용된 2개의 발현 강화 유전자좌는, 서열번호 1에 대해 실질적 상동성을 갖는 뉴클레오티드 서열을 포함하는 유전자좌, 서열번호 2에 대해 실질적 상동성을 갖는 뉴클레오티드 서열을 포함하는 유전자좌, 및 서열번호 3에 대해 실질적 상동성을 갖는 뉴클레오티드 서열을 포함하는 유전자좌로 이루어지는 군으로부터 선택된다. 이들 유전자좌는 서열에 대한 작동 가능한 연결 내에 (즉, 서열 내 또는 서열에 인접한 범위 내에) 통합된 유전자의 발현 강화를 제공할 뿐 아니라, 게놈 내의 다른 서열과 비교하여 더 큰 재조합 효율 및 강화된 통합 안정성을 나타낸다.

[0133] 서열번호 1, 서열번호 2 및 서열번호 3은 CHO 세포로부터 동정되었다. (예컨대 인간 또는 마우스와 같은) 다른 포유류 종은 동정된 발현 강화 영역에 대해 제한된 상동성을 갖는 것으로 확인되었지만, 상동 서열은 중국 햄스터(*Cricetulus griseus*,)의 다른 조직 유형 또는 다른 상동의 종으로부터 유래된 세포주에서 발견될 수 있고, 당업계에 널리 공지된 기법에 의해 분리될 수 있다. 예컨대, 중간 혼성화 또는 PCR-기반 기법에 의해 다른 상동 서열을 동정할 수 있다. 또한, 당업계에 널리 공지된 부위 지향적 또는 무작위 돌연변이 유발 기술에 의해 서열번호 1, 서열번호 2 또는 서열번호 3에 제시된 뉴클레오티드 서열을 변경할 수 있다. 이어서, 생성된 서열 변이체는 발현 강화 활성에 대해 시험될 수 있다. 발현 강화 활성을 갖고, 서열번호 1, 서열번호 2 또는 서열번호 3과 핵산 동일성이 적어도 약 90% 동일한 DNA는 통상적인 실험에 의해 분리되며, 발현 강화 활성을 나타낼 것으로 기대된다.

[0134] 통합 부위, 하나 이상의 외인성 핵산이 삽입된 부위 또는 뉴클레오티드 위치는 발현 강화 서열(예컨대, 서열번호 1, 서열번호 2, 서열번호 3) 중 어느 하나의 내에 있거나 이에 인접한 임의의 위치일 수 있다. 관심 유전자좌 내 또는 이에 인접한 특정 염색체 위치가 통합된 외인성 유전자의 안정한 통합 및 효율적인 전사를 지원하는지 여부는 당업계에 널리 공지된 표준 절차에 따라, 예를 들어 미국 특허 제8,389,239호 및 미국 특허 출원 일련번호 제14,919,300호에 기술된 바와 같이 결정될 수 있다.

[0135] 본원에서 고려되는 통합 부위는 발현 강화 서열 내, 또는 서열에 근접한 범위 내, 예를 들어, 염색체 DNA 상의 발현 강화 서열의 위치에 대해 약 1 kb 미만, 500 염기쌍(bp), 250 bp, 100 bp, 50 bp, 25 bp, 10 bp, 또는 약

5 bp 미만의 상류(5') 또는 하류(3')에 위치한다. 또 다른 일부 구현예에서, 사용된 통합 부위는 염색체 DNA 상의 발현 강화 서열의 위치에 대해 약 1000, 2500, 또는 5000 이상의 염기쌍 상류(5') 또는 하류(3')에 위치한다.

[0136] 큰 게놈 영역, 예컨대 지지체/기질 부착 영역이 염색체 DNA의 효율적인 복제 및 전사를 위해 사용되는 것으로 당 업계에서 이해된다. 지지체-부착 영역(SAR) 또는 기질-결합 또는 기질 부착 영역(MAR)으로도 일컬어지는 것으로 공지된 지지체/기질 부착 영역(S/MAR)은 핵 기질이 부착되는 진핵생물 게놈 DNA 영역이다. 임의의 하나의 이론에 구속됨이 없이, S/MAR은 전형적으로 비-암호화 영역에 도표화(mapping)되고, 주어진 전사 영역(예컨대, 염색질 도메인)을 이의 인접으로부터 분리하고, 전사를 가능하게 하는 인자, 예컨대 DNAses 또는 중합효소 인식 부위를 위한 기구 및/또는 결합을 위한 플랫폼이 또한 제공된다. 몇몇 S/MAR은 길이가 약 14 내지 20 kb인 것으로 특성분석 되었다(Klar 등, 2005, *Gene* 364:79-89). 이와 같이, 발현 강화 유전자좌에서의 (예컨대, 서열번호 1 또는 서열번호 2, 또는 서열번호 3 내에서 또는 그 근처에서의) 유전자 통합이 강화된 발현을 제공할 것으로 기대된다. 일부 구현예에서, 발현 강화 유전자좌 내의 특이적 부위에서 통합된 이중 특이적 항원 결합 단백질을 암호화하는 외인성 핵산 서열을 포함하는 숙주 세포는 높은 특이적 생산성을 나타낸다. 다른 구현예에서, 이중 특이적 항원 결합 단백질을 암호화하는 숙주 세포는 적어도 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 또는 30 피코그램/세포/일(pcd)의 특이적 생산성을 갖는다.

[0137] 일부 구현예에서, 외인성 핵산은 서열번호 1의 뉴클레오티드 서열을 포함하는 유전자좌 내의 부위에서 통합된다. 특이적 구현예에서, 통합 부위는 서열번호 1의 뉴클레오티드 서열 내, 또는 이에 대한 근접한 범위 내에 있다. 특정 구현예에서, 통합 부위는 10-13,515; 20-12,020; 1,020-11,020; 2,020-10,020; 3,020-9,020; 4,020-8,020; 5,020-7,020; 6,020-6,920; 6,120-6,820; 6,220-6,720; 6,320-6,620; 6,420-6,520; 6,460-6,500; 6,470-6,490; 및 6,475-6,485로 넘버링 된 뉴클레오티드의 결핍 위치들로부터 선택된 서열번호 1 내의 위치에 있다. 다른 구현예에서, 통합 부위는 서열번호 1의 뉴클레오티드 5,000-7,400, 5,000-6,500, 6,400-7,400 및 서열번호 1의 뉴클레오티드 6,400-6,500으로 이루어지는 군으로부터 선택되는 서열 내에 있다. 특이적 구현예에서, 서열번호 1의 뉴클레오티드 6471 내지 6473의 "act" 삼중항 전, 후, 또는 내의 통합 부위.

[0138] 일부 구현예에서, 외인성 핵산은 서열번호 2 또는 서열번호 3의 뉴클레오티드 서열을 포함하는 유전자좌 내의 부위에서 통합된다. 특이적 구현예에서, 통합 부위는 서열 번호 2의 뉴클레오티드 서열 내, 또는 이에 대한 근접한 범위 내에 있다. 특이적 구현예에서, 통합 부위는 서열 번호 3의 뉴클레오티드 서열 내, 또는 이에 대한 근접한 범위 내에 있다. 일부 구현예에서, 통합 부위는 서열번호 3의 뉴클레오티드 1990-1991, 1991-1992, 1992-1993, 1993-1994, 1995-1996, 1996-1997, 1997-1998, 1999-2000, 2001-2002, 2002-2003, 2003-2004, 2004-2005, 2005-2006, 2006-2007, 2007-2008, 2008-2009, 2009-2010, 2010-2011, 2011-2012, 2012-2013, 2013-2014, 2014-2015, 2015-2016, 2016-2017, 2017-2018, 2018-2019, 2019-2020, 2020-2021 또는 2021-2022 내에 있다. 특이적 구현예에서, 통합은 서열번호 3의 뉴클레오티드 2001~2022에 있거나 그 내에 있다. 일부 구현예에서, 외인성 핵산은 서열번호 3의 뉴클레오티드 2001~2002 또는 뉴클레오티드 2021~2022 뉴클레오티드에서 또는 그 안에 삽입되고, 삽입의 결과 서열번호 3의 뉴클레오티드 2002-2021이 결실된다.

[0139] **발현 강화 유전자좌 내로의 부위 특이적 통합(Site-Specific Integration Into An Expression-Enhancing Locus)**

[0140] 발현 강화 유전자좌 내에, 즉 본원에 기술된 바와 같은 발현 강화 유전자좌 내의 한 부위 내에 하나 이상의 외인성 핵산을 부위 특이적 방식으로 통합시키는 것은, 예를 들어, 상동 재조합에 의한 방식 및 당업계에서 기술된 재조합 효소 매개 카세트 교환에 의한 방식을 비롯한 여러 가지 방식으로 달성될 수 있다 (예: 미국 특허 제 8,389,239호 및 본원에 기술된 종래 기술 참조).

[0141] 일부 구현예에서, 하나 이상의 외인성 핵산 또는 관심 유전자의 통합에 편리한 발현 강화 유전자좌 내에 적어도 2개, 즉 2개 이상의 상이한 재조합 효소 인식 서열(RRS)을 함유하는 세포가 제공된다. 이러한 세포는, 아래 본원 및 당업계에 기술된 바와 같은 상동 재조합을 포함하는 다양한 수단, 예를 들어 미국 특허 제 8,389,239호 및 그에 개시된 기술에 의해 2개 이상의 RRS를 함유하는 외인성 핵산 서열을 원하는 유전자좌 내에 도입함으로써 획득될 수 있다.

[0142] 특이적 구현예에서, 다수의 외인성 핵산의 통합에 편리한 발현 강화 유전자좌 내에 3개 이상의 재조합 효소 인식 서열(RRS)을 함유하는 세포가 제공된다. 특정 구현예에서, 예를 들어 2개의 별개 외인성 핵산의 통합을 매개할 수 있는 발현 강화 유전자좌 내에 3개의 상이한 재조합 효소 인식 서열(RRS)을 함유하는 세포로서, 게놈 내의 5' RRS 및 중간 RRS는 통합 대상 제1 외인성 핵산의 측면에 위치하는 5' RRS 및 3' RRS와 일치하고, 게놈 내

의 중간 RRS 및 3' RRS는 통합 대상 제2 외인성 핵산의 측면에 위치하는 5' RRS 및 3' RRS와 일치하는 세포가 제공된다.

[0143] 적합한 RRS는 *LoxP*, *Lox511*, *Lox5171*, *Lox2272*, *Lox2372*, *Loxm2*, *Lox-FAS*, *Lox71*, *Lox66* 및 이들의 돌연변이를 포함하는 군으로부터 선택될 수 있으며, 부위 특이적 재조합 효소는 Cre 재조합 효소이거나 이의 유도체가 재조합 효소 매개 카세트 교환(RMCE)을 달성하는 데 사용된다. 다른 실시예에서, 적합한 RRS는 FRT, F3, F5, FRT 돌연변이-10, FRT 돌연변이+10 및 이들의 돌연변이를 포함하는 군으로부터 선택될 수 있으며, 이러한 시나리오에서는, 부위 특이적 재조합 효소 Flp 재조합 효소 또는 이의 유도체가 RMCE를 달성하는 데 사용된다. 또 다른 실시예에서, RRS는 attB, attP 및 이들의 돌연변이를 포함하는 군으로부터 선택될 수 있으며, 이 경우에는, 부위 특이적 재조합 효소 phiC31 인테그라아제 또는 이의 유도체가 RMCE를 달성하는 데 사용된다.

[0144] 다른 구현예에서, 원시 세포가 상동 재조합 기술에 의해 변형되어 하나 이상의 외인성 핵산을 함유하는 핵산 서열을 발현 강화 유전자좌 내의 특이적 부위에 통합시킨다.

[0145] 상동 재조합의 경우, 상동성 폴리뉴클레오티드 분자(즉, 상동성 아암)가 정렬되어 이들의 서열 구간을 교환한다. 전이유전자가 상동성 게놈 서열에 의해 측부에 위치할 경우, 이러한 교환 동안 전이 유전자가 도입될 수 있다. 하나의 실시예에서, 재조합 효소 인식 부위는 상동 재조합에 의해 통합 부위에서 숙주 세포 게놈 내에 도입될 수 있다. 다른 실시예에서, 하나 이상의 외인성 관심 핵산, 예를 들어, 각각 HCF 또는 LCF(예컨대, 가변 영역)을 암호화하는 하나 이상의 핵산을 함유하는 핵산 서열로서, 표적 유전자좌에서의 서열과 상동인 서열("상동성 아암")이 그 측면에 위치하는 핵산이 숙주 게놈 내에 삽입된다.

[0146] 진핵 세포에서 상동 재조합은 상기 통합 부위에서 염색체 DNA에 절단을 도입함으로써 촉진될 수 있다. 이는 통합 특정 부위에 특정 뉴클레아제를 표적화하여 수행될 수 있다. 상기 표적 유전자좌에서 DNA 서열을 인식하는 DNA-결합 단백질은 당해 기술 분야에 알려져 있다. 유전자 표적화 벡터는 또한 상동 재조합을 용이하게 하는데 사용된다.

[0147] 상동 재조합을 달성하기 위한 유전자 표적화 벡터의 작제 및 뉴클레아제의 선택은 본 발명이 속하는 당업자의 기술 범위 내에 있다. 일부 실시예에서, 모듈 구조를 가지며 개별 징크 핑거 도메인(zinc finger domains)을 함유하는 징크 핑거 뉴클레아제(ZFN)는 표적 서열(예를 들어, 표적화 통합 부위) 내 특정 3-뉴클레오티드 서열을 인식한다. 일부 구현예에서, 다수의 표적 서열을 표적화하는 개별 징크 핑거 도메인의 조합과 함께 ZFN이 사용될 수 있다. 전사 촉진제 유사(TAL) 이펙터 뉴클레아제(TALEN) 또한 부위 특이적 게놈 편집을 위해 사용될 수 있다. TAL 이펙터 단백질 DNA-결합 도메인은 일반적으로 FokI과 같은 제한 뉴클레아제의 비-특이적 절단 도메인과 조합하여 사용된다. 일부 구현예에서, TAL 이펙터 단백질 DNA-결합 도메인 및 제한 뉴클레아제 절단 도메인을 포함하는 융합 단백질은 본 발명의 유전자좌 내 표적 서열에서 DNA를 인식하고 절단하는데 사용된다(Boch J 등, 2009 *Science* 326:1509-1512). RNA 유전자가위(RNA-guided endonucleases; RGEN)는 세균의 적응 면역 기구로부터 개발된 도구를 조작하는 프로그램할 수 있는 게놈이다. 이 시스템-간헐적으로 반복되는 회문 구조 염기 서열 집합체(the clustered regularly interspaced short palindromic repeats; CRISPR)/CRISPR-관련(Cas) 면역 반응-에서 두 개의 RNA로 복합체를 형성할 때 상기 단백질 Cas9은 서열-특이적 엔도뉴클레아제를 형성하며 이들 RNA중 하나는 표적 선택을 유도한다. RGEN은 구성성분(Cas9 및 tracrRNA) 및 표적-특이적 CRISPR RNA(crRNA)로 이루어진다. DNA 표적 절단 효율 및 절단 부위의 위치 모두는 프로토스페이스 인접 모티프(proto-spacer adjacent motif; PAM)의 위치, 표적 인식에 대한 추가 요건에 따라 다양하다(원고 M113.539726로서 2014년 3월 14일자 온라인 공개된 Chen, H. 등의 *J. Biol. Chem.*). 특정 표적 유전자좌에 대해 고유한 서열(예컨대, 서열번호 1, 서열번호 2, 또는 서열번호 3)은, 16~17 염기쌍 일치로 잠재적인 표적 외 부위를 밝혀낼 수 있는 CHO 게놈에 이들 서열 중 많은 서열을 정렬함으로써 동정될 수 있다.

[0148] 일부 구현예에서, 관심 핵산(예: 하나 이상의 선별 표지 유전자의 측면에 선택적으로 위치하는 하나 이상의 RRS를 함유하는 핵산, 또는 각각 (가변 영역과 같은) HCF 또는 LCF를 암호화하고, 5' 및 3' 상동 아암이 측면에 위치한 하나 이상의 외인성 핵산을 함유하는 핵산)을 가지는 표적화 벡터가 하나 이상의 추가 벡터 또는 mRNA를 갖는 세포 내에 도입된다. 하나의 구현예에서, 하나 이상의 추가 벡터 또는 mRNA는 징크 핑거 뉴클레아제(ZFN), ZFN 이량체, 전사 촉진제 유사 이펙터 뉴클레아제(TALEN), TAL 이펙터 도메인 융합 단백질 및 RNA-유도된 DNA 엔도뉴클레아제를 포함하되 이들로 한정되지 않는 부위 특이적 뉴클레아제를 암호화하는 뉴클레오티드 서열을 함유한다. 특정 구현예에서, 하나 이상의 벡터 또는 mRNA는 가이드 RNA, tracrRNA 및 Cas 효소를 암호화하는 뉴클레오티드 서열을 포함하는 제1벡터, 및 공여자(외인성) 뉴클레오티드 서열을 포함하는 제2 벡터를 포함한다. 이러한 공여자 서열은 관심 유전자를 암호화하는 뉴클레오티드 서열, 또는 인식 서열, 또는 표적화된 삽입을 위

한 임의의 하나의 이러한 외인성 요소를 포함하는 유전자 카세트를 함유한다. mRNA가 사용되는 경우, mRNA는 당업자에게 공지된 통상의 형질감염 방법에 의해 세포 내로 형질감염될 수 있으며, 효소, 예컨대 전이효소 또는 엔도뉴클레아제를 암호화할 수 있다. 세포 내로 도입되는 mRNA는 일시적일 수 있고 계층 내로 통합되지 않지만, mRNA는 통합이 발생하기 위해 필수적이거나 유익한 외인성 핵산을 운반할 수 있다. 일부 경우에, mRNA는 부속 폴리뉴클레오티드의 부작용이 장기간 지속되는 임의의 위험을 없애기 위해 선택되는데, 이 경우, 핵산의 원하는 통합을 달성하기 위해서는 단기간의 발현만이 요구된다.

[0149] **부위 특이적 통합을 위한 벡터(Vectors For Site Specific Integration)**

[0150] 부위 특이적 통합을 통해 외인성 핵산을 2개의 발현 강화 유전자좌 내에 도입하기 위한 핵산 벡터가 본원에 제공된다. 적합한 벡터는, RMCE를 통한 통합을 위해 RRS가 측면에 위치한 외인성 핵산 서열을 함유하도록 설계된 벡터, 및 상동 재조합을 통한 통합을 위해 상동성 아암이 측면에 위치한 외인성 관심 핵산을 함유하도록 설계된 벡터를 포함한다.

[0151] 다양한 구현예에서, 벡터는 RMCE를 통한 부위 특이적 통합을 달성하기 위해 제공된다. 일부 구현예에서, 벡터는 2개의 표적 유전자좌 내에 다수의 핵산의 동시 통합을 달성하도록 설계된다. 순차적 통합과는 대조적으로, 동시 통합은 대규모 생산(제조)에 적합한 항원 결합 단백질 또는 다른 관심 다량체 단백질을 생산하는 원하는 클론의 효율적이고 신속한 단리를 가능하게 한다.

[0152] 일부 구현예에서, 세포에서 이중 특이적 항원 결합 단백질을 발현시키기 위한 벡터 세트가 제공된다.

[0153] 일부 구현예에서, 벡터 세트는 각각 5' RRS 및 3' RRS가 측면에 위치하는 핵산을 함유하는 2개의 "HCF 벡터"를 포함할 수 있는데, 상기 핵산은 HCF를 암호화하는 뉴클레오티드 서열을 포함하고, 2개의 HCF는 상이하다. 2개의 HCF 벡터 상의 RRS는 서로 상이하며, HCF를 암호화하는 뉴클레오티드 서열을 2개의 발현 강화 유전자좌에 통합하도록 설계된다. 또한, 벡터 세트는, HCF 벡터 중 하나 또는 2개의 HCF 벡터 모두에 포함될 수 있거나 (따라서, 동일한 LCF의 2개의 카피를 제공하거나), 대안적으로, 별도의 "LCF 벡터"로 제공될 수 있고 5' RRS 및 3' RRS가 측면에 위치할 수 있는 LCF를 암호화하는 뉴클레오티드 서열을 포함한다.

[0154] 일부 구현예에서, LCF 코딩 뉴클레오티드 서열은 HCF 벡터 중 하나에 포함되고, 상기 HCF 상에서 5' RRS와 3' RRS 사이에 위치한다. LCF 코딩 서열은 HCF 코딩 서열의 상류 또는 하류에 위치할 수 있다.

[0155] 일부 구현예에서, LCF 코딩 뉴클레오티드 서열은 HCF 벡터의 모두에 포함되고, 각각의 HCF 상에서 5' RRS와 3' RRS 사이에 위치한다. 마찬가지로, LCF 코딩 서열은 각각의 벡터에서 HCF 코딩 서열의 상류 또는 하류에 위치할 수 있다.

[0156] 일부 구현예에서, LCF 코딩 뉴클레오티드 서열은 별도의 벡터인 "LCF" 벡터로 제공되고, 5' RRS와 3' RRS가 측면에 위치하되, 2개의 RRS는 서로 상이하다. 벡터 세트에서의 RRS는, LCF 코딩 서열이 공통 RRS를 또한 함유하는 표적 유전자좌와 RMCE하는 동안공통 RRS를 통해 HCF 코딩 서열 중 하나와 "결합"될 수 있도록 설계될 수 있다. 예를 들어, LCF 벡터의 3' RRS는 하나의 HCF 벡터의 5' RRS와 동일할 수 있으며, RMCE를 통한 표적 유전자좌에서의 통합 후 LCF-HCF 배열을 생성시킨다. 다른 예에서, HCF 벡터의 3' RRS는 LCF 벡터의 5' RRS와 동일할 수 있으며, RMCE를 통한 표적 유전자좌에서의 통합 후 HCF-LCF 배열을 생성시킨다. 일부 구현예에서, 공통 RRS는 분리된 선택 가능 표지 포맷으로 설계될 수 있으며, 하나의 벡터에 포함된 선별 표지 유전자의 5' 부분의 3' 말단에 포함되고, 또 다른 벡터에 포함된 동일한 선별 표지 유전자의 나머지 3' 부분의 5' 말단에도 포함됨으로써, 표적 유전자좌 내에 "결합"되고 통합될 때, 적절히 통합된 핵산이 형질감염체가 편리하게 동정될 수 있도록 전체 선별 표지 유전자를 포함한다. 일부 구현예에서, 공통 RRS는 분리된 유전자 포맷으로 설계된다. 즉, 공통 RRS는, 하나의 벡터 상에서 유전자의 5' 부분의 일부로서 상기 유전자의 5' 부분의 3' 말단에 포함되거나 이렇게 분리된 유전자의 인트론 내에 포함되고, 분리된 유전자의 나머지 3' 부분의 일부로서 상기 유전자의 나머지 부분의 5' 단부에 포함되거나 이렇게 분리된 유전자의 인트론 내에 포함된다. 또 다른 구현예에서, 제1 벡터에서의 제2 또는 중간 RRS는 제2 또는 중간 RRS가 작동 가능하게 연결되는 (다른 하나의 벡터로부터는 분리됨) 선별 표지 유전자와 프로모터 사이에 있도록 설계되고, 제1 벡터에서의 제3 또는 중간 RRS는 프로모터의 3'에 있도록 설계되며, 제2 벡터에서의 제3 또는 중간 RRS는 선별 표지 유전자의 5'에 있도록 설계된다.

[0157] 일부 구현예에서, 벡터 세트는 LCF를 암호화하는 추가 뉴클레오티드 서열을 포함할 수 있다. 다시 말해, 벡터 세트는 2개의 HCF 벡터 및 2개의 LCF 암호화 뉴클레오티드 서열을 포함할 수 있다. 2개의 LCF 암호화 서열은 동일하거나 상이한 LCF를 암호화할 수 있다. 일부 구현예에서, 2개의 LCF 암호화 서열은 HCF 벡터 내에 각각 포함되어, HCF 코딩 서열과 LCF 코딩 서열을 각각 함유하는 2개의 벡터가 생성될 수 있다. 2개의 벡터는, 2개의 벡

터를 2개의 유전자좌 내에 표적화하기에 적합한 RRS를 갖도록 설계될 수 있다. 다른 구현예에서, 2개의 LCF 암호화 서열 중 하나는 HCF 벡터 내에 포함되어 상기 HCF 벡터 상에서 5' RRS와 3' RRS 사이에 위치하고, 다른 하나의 LCF 암호화 서열은 별도의 벡터 상에 제공된다 - 즉, 하나의 벡터가 (LCF-HCF 또는 HCF-LCF 배열 내, 또는 줄여서 "LCF/HCF 벡터"에서) LCF와 HC 모두, 즉 하나의 HCF 벡터와 하나의 LCF 벡터를 가진다. 이들 다른 구현예 중 일부에서, 벡터 RRS는 HCF 벡터 상의 HCF 코딩 서열과 LCF 벡터 상의 LCF 코딩 서열이 표적 유전자좌에서 RMCE를 통해 결합하는 것을 허용하도록 설계될 수 있다. 예를 들어, LCF 벡터의 3' RRS는 HCF 벡터의 5' RRS와 동일할 수 있고, 공통 RRS는 형질감염체의 선별과 동정을 용이하게 하기 위해 분리된 선별 표지 포맷 또는 분리된 인트론 포맷으로 설계될 수 있다. 또 다른 구현예에서, 2개의 LCF가 상이한 경우, 2개의 LCF 코딩 뉴클레오타이드 서열은 각각 별도의 벡터 상에 제공될 수 있다 - 즉, 벡터 세트는 2개의 HCF 벡터 및 2개의 LCF 벡터를 포함한다. RRS는 하나의 표적 유전자좌에서 하나의 LCF 코딩 서열이 하나의 HCF 코딩 서열과 적절한 "결합"되고, 제2 표적 유전자좌에서 다른 하나의 LCF 코딩 서열이 다른 하나의 HCF 코딩 서열과 적절한 결합되도록 설계될 수 있다. 도 1, 3 및 4는 벡터의 상이한 포맷 및 RRS/유전자좌 조합의 예시이며 한정하려는 것은 아니다. 각각의 주어진 벡터 시스템은 양성 인티그란트(원하는 클론)를 신속하고 편리하게 선별하기 위해 재조합 효소의 존재 하에 각각의 뉴클레오타이드 서열을 동시에 통합하기 위한 수단을 제공한다.

[0158] HCF 또는 LCF를 암호화하는 뉴클레오타이드 서열은 불변 영역으로부터 아미노산 또는 도메인(들)을 암호화하거나, 전체 불변 영역을 암호화할 수 있다. 특정 구현예에서, HCF 또는 LCF를 암호화하는 뉴클레오타이드 서열은 하나 이상의 불변 영역, 예컨대, CL, CH1, 힌지, CH2, CH3 또는 이들의 조합을 암호화할 수 있다. 특정 구현예에서, HCF 도메인을 암호화하는 뉴클레오타이드 서열은 CH3 도메인을 암호화할 수 있다. 예를 들어, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 CH3 도메인을 암호화할 수 있고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 CH3 도메인을 암호화할 수 있다. 제1 및 제2 CH3 도메인은 동일하거나, 적어도 하나의 아미노산이 상이하다. CH3 도메인 또는 불변 영역에서의 차이(예: 상이한 단백질 A 결합 특성으로 나타나거나 "노브 앤 홀(knob-and-hole)" 포맷인 차이)은 본원에 기술된 이중 특이적 항원 결합 단백질에 대해 임의의 포맷을 취할 수 있다. 임의의 아미노산 서열 차이와 별개로, 2개의 HCF 코딩 뉴클레오타이드 서열은 하나 이상의 뉴클레오타이드 서열이 코돈 최적화되었다는 것이 또한 상이하다.

[0159] 일부 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 가령 프로모터를 함유하는 전사 조절 서열에 독립적으로 및 작동 가능하게 연결된다. 일부 구현예에서, 2개의 HCF 함유 폴리펩티드의 전사를 유도하는 프로모터들은 동일하다. 일부 구현예에서, 2개의 HCF 함유 폴리펩티드의 전사를 유도하는 프로모터들뿐만 아니라 LCF 함유 폴리펩티드의 전사를 유도하는 프로모터(예: CMV 프로모터, 또는 본원에 기술된 임의의 다른 적합한 프로모터)도 모두 동일하다. 일부 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 유도 프로모터 또는 억제 프로모터에 독립적으로 및 작동 가능하게 연결된다. 유도 프로모터 또는 억제 프로모터는, 예를 들어, 생산 단계(유가 배양)에서만 생산이 이루어지도록 하고, 성장 단계(시드 트레인 배양) 동안에는 생산이 이루어지지 않도록 한다. 유도 프로모터 또는 억제 프로모터는 또한 하나 이상의 관심 유전자의 차별적 발현이 이루어지도록 한다. 일부 구현예에서, 각각의 HCF 및/또는 LCF 코딩 뉴클레오타이드 서열은 적어도 하나의 TetR 작동자(TetO) 또는 Arc 작동자(ArcO)의 상류에서 프로모터에 독립적으로 및 작동 가능하게 연결된다. 다른 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 CMV/TetO 또는 CMV/ArcO 하이브리드 프로모터에 독립적으로 및 작동 가능하게 연결된다. 하이브리드 프로모터(조절 융합 단백질로도 지칭됨)의 실시예는 2003년 12월 11일에 공개된 국제 공보 제W003101189A1호에서 (본원에 참조로서 통합됨) 확인할 수 있다.

[0160] 일부 구현예에서, 벡터 세트는, LCF 암호화 벡터나 HCF 암호화 벡터 중 하나에 포함되거나 별도 벡터로 제공될 수 있는 하나 이상의 RRS를 인식하는 재조합 효소를 암호화하는 뉴클레오타이드 서열을 포함한다.

[0161] 다양한 다른 구현예에서, 벡터는 상동 재조합을 통해 부위 특이적 통합을 달성하도록 제공된다.

[0162] 일부 구현예에서, 세포 내 2개의 유전자좌 내로의 부위 특이적 통합을 위한 5' 상동 아암과 3' 상동 아암이 측면에 위치한 외인성 핵산을 각각 함유하는 2개의 벡터를 포함하는 벡터 세트가 제공되며, 상기 2개의 벡터의 외인성 핵산은 함께 항원 결합 단백질을 암호화한다. 따라서, 하나의 벡터 상의 상동 아암은 2개의 유전자좌 중 하나 내로의 통합을 위해 설계되고, 다른 하나의 벡터 상의 상동 아암은 다른 하나의 유전자좌 내로의 통합을 위해 설계된다. 이들 구현예에서, 항원 결합 단백질은 단일 특이적이거나 이중 특이적일 수 있다.

[0163] 발현 강화 유전자좌 내의 서열과 상동인 서열을 선별하고, 선별된 서열을 표적 벡터에 상동 아암으로서 포함시키는 것은 당업자의 기술 범위 내에 있다. 일부 구현예에서, 벡터 또는 작제물은 제1 상동 아암 및 제2 상동 아암을 포함하며, 결합된 제1 및 제2 상동 아암은 유전자좌 내에서 내인성 서열을 치환하는 표적화 서열을 포함한다.

다. 또 다른 구현예에서, 제1 및 제2 상동 아암은 유전자와 내의 내인성 서열 내에 통합되거나 삽입되는 표적화 서열을 포함한다. 일부 구현예에서, 상동 아암은 서열번호 1, 서열번호 2, 또는 서열번호 3에 존재하는 뉴클레오티드 서열과 상동인 뉴클레오티드 서열을 포함한다. 특이적 구현예에서, 벡터는 서열번호 3의 뉴클레오티드 1001~2001에 상응하는 뉴클레오티드 서열을 갖는 5' 상동 아암, 및 서열번호 3의 뉴클레오티드 2022~3022에 상응하는 뉴클레오티드를 갖는 3' 상동 아암을 함유한다. 상동 아암, 예컨대 제1 상동 아암(5' 상동 아암으로도 불림) 및 제2 상동 아암(3' 상동 아암으로도 불림)은 유전자와 내에서의 표적화 서열에 상동이다. 상동 아암은, 적어도 1 kb, 또는 적어도 약 2 kb, 또는 적어도 약 3 kb, 또는 적어도 약 4 kb, 또는 적어도 5 kb, 또는 적어도 약 10 kb를 포함하는 유전자와 내에서의 영역 또는 표적화 서열을 5'에서 3'까지 확장시킬 수 있다. 다른 구현예에서, 제1 및 제2 상동 아암으로부터 선택된 표적화된 서열의 총 뉴클레오티드 수는 1 kb 이상, 또는 약 2 kb 이상, 또는 약 3 kb 이상, 또는 약 4 kb 이상, 또는 5 kb 이상, 또는 약 10 kb 이상이다. 일부 예에서, 5' 상동 아암과 3' 상동 아암(표적화 서열과 상동임) 사이의 거리는 적어도 5 bp, 10 bp, 20 bp, 30 bp, 40 bp, 50 bp, 60 bp, 70 bp, 80 bp, 90 bp, 100 bp, 200 bp, 300 bp, 400 bp, 500 bp, 600 bp, 700 bp, 800 bp, 900 bp, 또는 적어도 1 kb, 또는 적어도 약 2 kb, 또는 적어도 약 3 kb, 또는 적어도 약 4 kb, 또는 적어도 5 kb, 또는 적어도 약 10 kb를 포함한다. 서열번호 3의 뉴클레오티드 1001~2001 및 2022~3022가 5' 및 3' 상동 아암으로서 선택된 예에서, 2개의 상동성 아암 사이의 거리는 20 뉴클레오티드(서열번호 3의 뉴클레오티드 2002~2021에 상응함)일 수 있으며; 이러한 상동 아암은 서열번호 3을 포함하는 유전자와 내에서의 (예를 들어, 서열번호 3의 뉴클레오티드 1990~2021 또는 2002~2021 내에서의) 외인성 핵산 서열의 통합과, 서열번호 3의 뉴클레오티드 2002~2021의 동시 결실을 매개할 수 있다.

[0164] 발현 강화 유전자와 내로의 부위 특이적 통합을 위한 외인성 핵산을 도입하기 위한, 본원에 개시된 벡터는, 외인성 관심 핵산 및 암호화된 폴리펩티드의 발현을 유도하고, 외인성 관심 핵산이 성공적으로 통합된 세포의 선별하고 동정하기 위한 추가 유전자 및 서열을 포함할 수 있다. 이러한 추가 서열은, 예를 들어, 아래 본원에도 기술된 전사 및 번역 조절 서열, 선별 표지 유전자 등을 포함한다.

[0165] 조절 서열(Regulatory Sequences)

[0166] 발현 강화 유전자와 내에 외인성 핵산을 부위 특이적 방식으로 도입하기 위한, 본원에 개시된 벡터 및 부위 특이적 통합의 결과로서 획득된 세포는 외인성 관심 핵산 및 암호화된 폴리펩티드의 발현을 유도하기 위한 조절 서열을 포함할 수 있다. 조절 서열은 전사 프로모터, 증강자, 적합한 mRNA 리보솜 결합 부위를 암호화하는 서열, 및 전사 및 번역의 종결을 제어하는 서열을 포함한다. 전사 및 번역 조절 서열은 바이러스 공급원에 의해 제공될 수 있다. 예를 들어, 공통적으로 사용된 프로모터 및 증강자는 폴리오마(polyoma), 아데노바이러스(adenovirus) 2, 시미안 바이러스 40 (SV40), 마우스 또는 인간 사이토메갈로바이러스 (CMV), CMV 극초기 (CMV-IE) 또는 CMV 주요 IE (CMV-MIE) 프로모터뿐만 아니라 RSV, SV40 후기 프로모터, SL3-3, MMTV, 유비퀴틴 (Ubi), 유비퀴틴 C (Ubc) 및 HIV LTR 프로모터와 같은 바이러스로부터 유래한다. 이러한 제어 서열이 선택된 숙주 세포와 양립할 수 있다면, 바이러스 게놈 프로모터, 제어 및/또는 신호 서열이 발현을 유도하는 데 사용될 수 있다. 관심 단백질이 발현될 세포 유형에 따라, 비-바이러스 세포 프로모터(예컨대, b-글로빈 및 EF-1a 프로모터)가 사용될 수도 있다. SV40 바이러스 게놈, 예를 들어 SV40 기원, 초기 및 후기 프로모터, 증강자, 스플라이스 및 폴리아데닐화 부위로부터 유래된 DNA 서열을 사용해서 외인성 DNA 서열의 발현에 유용한 다른 유전 요소를 제공할 수 있다. 초기 및 후기 프로모터는 둘 다 SV40 바이러스 복제 기점을 또한 포함하는 단편으로서 SV40 바이러스에서 쉽게 획득되기 때문에 특히 유용하다(Fiers 등의 *Nature* 273:113, 1978). 더 작거나 더 큰 SV40 단편이 사용될 수도 있다. 통상적으로, Hind III 부위로부터 SV40 복제 기점에 위치한 BglI 부위를 향해 연장되는 약 250 bp의 서열이 포함된다. 유도성(예를 들어, 화학적 화합물, 보조 인자, 조절 단백질에 의해 유도되는) 및/또는 억제성(예를 들어, 화학적 화합물, 보조 인자, 조절 단백질에 의해 억제되는) 프로모터가 사용될 수 있는데, 이들은 항원 결합 단백질의 생산이 성장 단계(시드 트레인 배양) 동안이 아닌 생산 단계(유가 배양)에서만 일어나게 하거나, 상이한 유전자좌에서 항체 성분의 발현을 정밀하게 차별적으로 조절하는 데 특히 유용하다. 유도성 프로모터의 예에는 알코올 탈수소 효소 I 유전자 프로모터, 테트라시클린 반응성 프로모터 시스템, 글루코코르티코이드 수용체 프로모터, 에스트로겐 수용체 프로모터, 엑디손 수용체 프로모터, 메탈로티오네인 기반 프로모터 및 T7 폴리머라제 기반 프로모터가 포함된다. 억제성 프로모터의 예에는 CMV, 또는 적어도 하나의 TetR 오퍼레이터(TetO) 또는 아크 오퍼레이터(ArcO)에 작동 가능하게 연결된 다른 프로모터를 포함하는 하이브리드 프로모터(조절 융합 단백질로도 지칭됨)가 포함되며, 이는 2003년 12월 11일자로 공개된 국제 공보 제 WO03101189A1호(참조로서 본원에 통합됨)에 기술되어 있다. 바이시스트로닉 벡터(bicistronic vector)를 통해 다수의 전사물을 발현시키는 데 적합한 서열이 전술되었으며(Kim S. K. 및 Wold B. J., *Cell* 42:129, 1985), 본 발명에 사용될 수 있다. 단백질의 멀티시스트론(multicistronic) 발현을 위한 적합한 전략의 예에는 2A 펩티

드의 사용(Szymczak 등, *Expert Opin Biol Ther* 5: 627-638 (2005)) 및 내부 리보솜 개시 부위("IRES")가 포함되며, 둘 모두는 당업계에서 널리 공지되어 있다. 미국 특허 제4,634,665호(Axel 등) 및 미국 특허 제4,656,134호(Ringold 등)에 기재된 것들과 같은 다른 유형의 발현 벡터도 유용할 것이다.

[0167] **선별 표지(Selectable Markers)**

[0168] 외인성 핵산을 발현 강화 유전자와 내에 부위 특이적 방식으로 도입하기 위한, 본원에 개시된 벡터, 및 부위 특이적 통합의 결과로서 수득된 세포는 하나 이상의 선별 표지 유전자를 포함할 수 있다.

[0169] 일부 구현예에서, 선별 표지 유전자는, 예를 들어, Kaufman, R. J. (1988) *Meth Enzymology* 185:537의 표 1에 기술된 것들과 같은 약물 내성, DHFR-MTX 내성을 참조하며, DHFR-MTX 내성, P-당단백질 및 다중 약물 내성(MDR) - 다양한 친유성 세포독성 제제(예: 아드리아마이신, 콜히친, 빈크리스틴), 및 아데노신 탈아미노효소(ADA)-Xyl-A 또는 아데노신 및 2'-디옥시코포르마이신을 포함한다. 다른 우성 선별 표지는 미생물에 의해 유도된 항생제 내성 유전자, 예컨대 네오마이신, 카나마이신 또는 하이그로마이신 내성을 포함한다. 포유류 숙주에 대해 여러 가지 적절한 선별 시스템이 존재한다(전술된 Sambrook, pgs 16.9-16.15). 2개의 우성 선별 표지를 사용하는 동시-형질감염 프로토콜이 또한 기술되어 있다(Okayama 및 Berg, *Mol. Cell Biol* 5:1136, 1985).

[0170] 다른 구현예에서, 선별 표지 유전자는, 경우에 따라 성공적으로 삽입 및/또는 치환되었거나 성공적으로 삽입 및/또는 대체되지 않은 유전자 카세트의 인식을 위한 검출 가능한 신호를 제공하거나 생성할 수 있는 폴리펩티드를 암호화한다. 적합한 예에는 형광 표지 또는 단백질, 특히 검출 가능한 신호를 생성하는 화학 반응을 촉매하는 효소가 포함된다. 디스코스마 산호(DsRed), 녹색 형광 단백질(GFP), 향상된 녹색 형광 단백질(eGFP), 시안 형광 단백질(CFP), 향상된 시안 형광 단백질(eCFP), 황색 형광 단백질(YFP), 향상된 황색 형광 단백질(eYFP) 및 원-적색 형광 단백질(예컨대, mKate, mKate2, mPlum, mRaspberry 또는 E2- crimson)을 포함하되 이들로 한정되지 않는 형광 표지의 예가 당업계에 널리 공지되어 있다. 예를 들어, Nagai, T., 등의 2002 *Nature Biotechnology* 20:87-90; Heim, R. 등의 23 February 1995 *Nature* 373:663-664; 및 Strack, R.L. 등의 2009 *Biochemistry* 48:8279-81을 또한 참조한다.

[0171] **항원 결합 단백질 제작 시스템(Systems for Making Antigen-Binding Proteins)**

[0172] 추가적인 양태에서, 본 개시는, 세포(예: CHO 세포)를 하나 이상의 벡터와 재조합하는 것을 포함하는 시스템으로서, 단일 특이적 단백질 또는 이중 특이적 단백질 중 하나인 항원 결합 단백질을 함께 암호화하는 두 개의 발현 강화 유전자와 외인성 핵산 내에 통합된 세포를 만드는데 사용될 수 있는 시스템을 제공한다. 상기 시스템은, 예를 들어, 키트의 형태로 제공될 수 있다.

[0173] 일부 구현예에서, 효율적인 벡터가 구성될 수 있고 RMCE를 통해 다수의 외인성 핵산이 2개의 발현 강화 유전자와 내에 동시에 통합될 수 있도록 시스템이 설계된다. 동시 통합은 원하는 클론이 신속하게 단리될 수 있게 하며, 2개의 발현 강화 유전자와의 사용 또한 단백질 생산에 적합한 안정한 세포주(예: 상업적으로 이용 가능한 세포주)의 생성을 위해 중요하다.

[0174] 본원에 제공된 시스템에는 세포 및 벡터 세트가 포함된다. 세포는 2개의 발현 강화 유전자와 각각에 통합된 RRS 쌍(5' RRS 및 3' RRS)을 함유한다. 일부 구현예에서, 외인성 핵산은 각각의 유전자좌에서 5' RRS와 3' RRS 사이에 존재하고, 예를 들어, 하나 이상의 선별 표지 유전자를 포함할 수 있다. 벡터 세트는 적어도 2개의 벡터를 포함하는데, 각각의 벡터는 HCF 또는 LCF를 암호화하는 핵산 서열의 측면에 위치하는 RRS 쌍(5' RRS 및 3' RRS)을 함유하고, 2개의 벡터 중 하나의 뉴클레오티드 서열은 HCF를 암호화하거나(HCF 벡터), 2개의 벡터 중 다른 하나의 뉴클레오티드 서열은 LCF를 암호화하되(LCF 벡터), HCF와 LCF는 항원 결합 단백질의 영역이다. 각각의 RRS 쌍 내의 5' RRS와 3' RRS는 상이하고, 시스템에서의 RRS는 벡터가 세포 내에 도입될 때, 벡터의 HCF 또는 LCF 암호화 뉴클레오티드 서열이 항원 결합 단백질을 발현하도록 RRS의 의해 매개되는 RMCE를 통해 2개의 발현 강화 유전자와 내에 통합되도록 설계된다. 항원 결합 단백질, 벡터의 수, HCF 또는 LCF 코딩 서열의 배치, 및 RRS 간의 관계 여부에 따라 설계는 달라질 수 있다.

[0175] 일부 구현예에서, 시스템은 2개의 발현 강화 유전자와 내로의 통합 및 단일 특이적 항원 결합 단백질의 발현을 위해 설계된다. 일부 구현예에서, 2개의 벡터(즉 HCF 벡터 및 LCF 벡터) 중 하나에서의 5' RRS와 3' RRS는 2개의 유전자좌 중 하나에서의 5' RRS 및 3' RRS와 각각 동일하고, 다른 하나의 벡터에서의 5' RRS와 3' RRS는 다른 하나의 유전자좌에서의 5' RRS 및 3' RRS와 각각 동일하며, 본질적으로 HCF 및 LCF 핵산을 2개의 유전자좌에 각각 개별적으로 표적화한다. 다른 구현예에서, HCF 코딩 서열 및 LCF 코딩 서열은 별개의 벡터 상에 있지만 2개의 유전자좌와 각각에 결합되어 통합되도록 설계된다. 이들 구현예에 따르면, 제1 유전자좌에서의 5' RRS와 3'

RRS는 제2 유전자좌에서의 5' RRS 및 3' RRS와 동일하고, 각각의 유전자좌 또한 5' RRS와 3' RRS 사이에 추가 RRS("중간 RRS")를 함유한다. 또한, 제1 및 제2 벡터에서의 5' RRS는 제1 및 제2 유전자좌에서의 5' RRS와 동일하고, 상기 제1 벡터에서의 3' RRS는 제2 벡터에서의 5' RRS 및 제1 및 제2 유전자좌에서의 중간 RRS와 동일하며, 제2 벡터에서의 3' RRS는 2개의 유전자 모두에서의 3' RRS와 동일하다. 벡터는 분리된 프로모터와 선별 표지 포맷(하나의 벡터 상의 프로모터 및 프로모터가 또 다른 벡터 상에서 작동 가능하게 연결되는 선별 표지)을 가지도록 설계된다. 벡터는 적절한 통합과 함께 형질감염체의 선별을 용이하게 하기 위해, 전술한 바와 같이 분리된 선별 표지 포맷 또는 분리된 인트론 포맷을 가지도록 설계될 수 있다. 또한, 시스템은 통합 후 LCF 코딩 서열과 HCF 코딩 서열의 상이한 상대 위치를 허용하도록 설계될 수 있다. 일부 구현예에서, 시스템은 HCF 코딩 서열의 상류에 통합된 LCF 코딩 서열을 갖도록 설계된다. 다른 구현예에서, 시스템은 HCF 코딩 서열의 상류에 통합된 LCF 코딩 서열을 갖도록 설계된다.

[0176] 일부 구현예에서, 시스템은 2개의 발현 강화 유전자좌 내로의 통합 및 이중 특이적 항원 결합 단백질의 발현을 위해 설계된다.

[0177] 일부 구현예에서, (제1 HCF를 암호화하는) HCF 벡터 및 (제1 LCF를 암호화하는) LCF 벡터에 추가하여, 시스템은 제1 HCF와 상이한 제2 HCF를 암호화하는 뉴클레오타이드 서열을 또한 포함한다. 제2 HCF를 암호화하는 뉴클레오타이드 서열은, 예를 들어, LCF 벡터에 포함되거나 별도의 벡터(즉, 제2 HCF 벡터)에 포함될 수 있다. 일부 구현예에서, 제2 HCF 코딩 서열은 LCF 벡터 상의 5' RRS와 3' RRS 사이의 LCF 벡터에 포함되며, 이 경우, 시스템은 HCF 벡터 및 LCF/HCF 벡터를 포함한다. 시스템은, 특히 RRS는 HCF 코딩 서열을 2개의 유전자좌 중 하나의 유전자좌 내에 통합하고, HCF와 LCF 모두를 암호화하는 서열을 다른 하나의 유전자좌 내에 통합하도록 설계될 수 있다. 다른 구현예에서, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 5' RRS와 3' RRS가 측면에 위치한 별도의 벡터 상에 있으며, 이 경우, 시스템은 2개의 HCF 벡터 및 1개의 LCF 벡터를 포함한다. 이들 다른 구현예에서, LCF 코딩 서열이 RMCE를 통해 HCF 코딩 서열 중 하나와 공통 RRS(상기 유전자좌에서 5' RRS와 3' RRS 사이의 2개의 유전자좌 중 하나에도 존재함)를 통해 "결합"될 수 있고, 다른 하나의 HCF 코딩 서열은 2개의 유전자좌 중 다른 하나에 통합되도록 시스템에서의 RRS가 설계될 수 있다. 예를 들어, LCF 벡터의 3' RRS는 하나의 HCF 벡터의 5' RRS 및 2개의 유전자좌 중 하나의 중간 RRS와도 동일할 수 있는데 - 이러한 설계는 중간 RRS를 갖는 유전자좌 내로의 통합 후 LCF-HCF 배치를 생성할 것이다. 또 다른 구현예에서, HCF 벡터의 3' RRS는 LCF 벡터의 5' RRS 및 2개의 유전자좌 중 하나의 중간 RRS와 동일할 수 있는데, 이는 중간 RRS를 갖는 유전자좌 내로의 통합 후 HCF-LCF 배치를 생성할 것이다. 일부 구현예에서, 공통 RRS는 상기 본원에 기술된 바와 같이 분리된 선별 표지 포맷 또는 분리된 인트론 포맷으로 설계된다.

[0178] 일부 구현예에서, 시스템은 (제1 HCF를 암호화하는) HCF 벡터 및 (제1 LCF를 암호화하는) LCF 벡터, 및 제1 HCF와 상이한 제2 HCF를 암호화하는 뉴클레오타이드 서열에 추가하여 제2 LCF를 암호화하는 뉴클레오타이드 서열을 또한 포함한다. 즉, 시스템은 HCF를 암호화하는 2개의 코딩 서열 및 LCF를 암호화하는 2개의 코딩 서열인 4개의 별도 코딩 서열을 포함한다. 2개의 LCF는 동일하거나 상이할 수 있다. 4개의 코딩 서열은 상이한 설계로 벡터 내에 배치될 수 있다. 일부 구현예에서, 4개의 서열은 2개의 벡터 내에 배치된다. 벡터 중 하나에서 HCF의 상류 또는 하류에 LCF가 있는 LCF/HCF 및 LCF/HCF. 시스템(RRS)은, 하나의 벡터 서열이 하나의 유전자좌 내에 통합되고 다른 하나의 벡터 서열이 다른 하나의 유전자좌 내에 통합되도록 설계될 수 있다. 일부 구현예에서, 4개의 서열은 3개의 벡터 내에 배치된다. (HCF의 상류 또는 하류에 LCF가 있는) LCF, HCF, 및 LCF/HCF. 시스템에서의 RRS는, LCF/HCF 벡터에서의 서열이 하나의 유전자좌 내에 통합되고, LCF 벡터에서의 LCF 코딩 서열 및 HCF 벡터에서의 HCF 코딩 서열이 LCF 벡터, HCF 벡터 및 다른 하나의 유전자좌에 의해 공유된 공통 RRS를 사용하여 상기 다른 하나의 유전자좌 내에 통합되도록 설계될 수 있다. 마찬가지로, 공통 RRS는 분리된 표지 또는 분리된 인트론 포맷으로 설계될 수 있다. 일부 구현예에서, 4개의 서열은 4개의 벡터 내에 배치된다. LCF, HCF, LCF, 및 HCF. 시스템에서의 RRS는, LCF 벡터 중 하나에서의 LCF 코딩 서열 및 HCF 벡터 중 하나에서의 HCF 코딩 서열이 공통 RRS를 사용해 하나의 유전자좌 내에 결합되어 통합되고, LCF 벡터 중 다른 하나에서의 LCF 코딩 서열 및 HCF 벡터 중 다른 하나에서의 HCF 코딩 서열이 공통 RRS를 사용해 하나의 유전자좌 내에 결합되어 통합되도록 설계될 수 있다.

[0179] 본원에 제공된 시스템의 다양한 구현예에서, HCF 또는 LCF를 암호화하는 뉴클레오타이드 서열은 아미노산, 예를 들어, 불변 영역으로부터 아미노산 또는 도메인(들)을 암호화하거나, 전체 불변 영역을 암호화할 수 있다. 특이적 구현예에서, HCF 또는 LCF를 암호화하는 뉴클레오타이드 서열은 하나 이상의 불변 영역, 예컨대, CL, CH1, CH2, CH3 또는 이들의 조합을 암호화할 수 있다. 특정 구현예에서, HCF 도메인을 암호화하는 뉴클레오타이드 서열은 CH3 도메인을 암호화할 수 있다. 예를 들어, 제1 HCF를 암호화하는 뉴클레오타이드 서열은 제1 CH3 도메인을

암호화할 수 있고, 제2 HCF를 암호화하는 뉴클레오타이드 서열은 제2 CH3 도메인을 암호화할 수 있다. 제1 및 제2 CH3 도메인은 동일하거나, 적어도 하나의 아미노산이 상이하다. CH3 도메인 또는 불변 영역에서의 차이(예: 상이한 단백질 A 결합 특성으로 나타나거나 "노브 앤 홀(knob-and-hole)" 포맷인 차이)은 본원에 기술된 이중 특이적 항원 결합 단백질에 대해 임의의 포맷을 취할 수 있다. 임의의 아미노산 서열 차이와 별개로, 2개의 HCF 코딩 뉴클레오타이드 서열은 하나 이상의 뉴클레오타이드 서열이 코돈 최적화되었다는 것이 또한 상이하다.

[0180] 본원에 제공된 시스템의 다양한 구현예에서, 각각의 HCF 또는 LCF 코딩 뉴클레오타이드 서열은 가령 프로모터를 포함하는 전사 조절 서열에 독립적으로 및 작동 가능하게 연결된다. 일부 구현예에서, 2개의 HCF 함유 폴리펩티드의 전사를 유도하는 프로모터들은 동일하다. 일부 구현예에서, 2개의 HCF 함유 폴리펩티드의 전사를 유도하는 프로모터들뿐만 아니라 LCF 함유 폴리펩티드의 전사를 유도하는 프로모터(예: CMV 프로모터, 유도성 프로모터, 억제성 프로모터, 또는 본원에 기술된 임의의 다른 적합한 프로모터)도 모두 동일하다.

[0181] 일부 구현예에서, 본 시스템은, 가변 영역 코딩 벡터 중 하나에 포함되거나 별도 벡터로 제공될 수 있는 하나 이상의 RRS를 인식하는 재조합 효소를 암호화하는 뉴클레오타이드 서열을 추가로 포함한다.

[0182] 본원에 개시된 시스템은 벡터의 효율적인 구성 및 바람직한 클론의 신속한 단리를 허용하도록 설계되며, 2개의 발현 강화 유전자좌의 사용 또한 안정한 세포주의 생성을 위해서 중요하다. 일부 구현예에서, 시스템은 음성 선별을 사용해 의도된 부위 특이적 통합(예: RMCE 이후에 제거되는 숙주 게놈 내의 하나 이상의 형광 표지 유전자에 기인한 형광의 결핍)을 갖는 형질전환체를 동정하도록 설계된다. 한 번의 음성 선별은 불과 2주일이 걸리지만, 의도된 재조합으로 클론을 단리하는 효율은 제한적일 수 있다(약 1%). 음성 선별이 통합된 핵산(들)에 의해 제공된 새로운 선별 표지(예컨대, 새로운 형광 표지, 또는 예를 들어 분리된 포맷의 약물 또는 항생제 내성)에 기초한 양성 선별과 조합되는 경우, 의도된 재조합으로 클론을 단리하는 효율은 (약 40% 내지 약 80%까지) 상당히 개선될 수 있다. 시스템은 추가적인 성분, 시약, 또는 정보, 예를 들어, 시스템의 벡터(들)를 형질감염에 의해 시스템의 세포 내에 도입하기 위한 프로토콜을 포함할 수 있다. 비제한적인 형질감염 방법은 화학-기반 형질감염 방법을 포함하고, 이는 리포좀; 나노입자; 칼슘 포스페이트(Graham 등 (1973) *Virology* 52 (2): 456-67, Bacchetti 등 (1977) *Proc Natl Acad Sci USA* 74 (4): 1590-4 및, Kriegler, M (1991) *Transfer and Expression: A Laboratory Manual*. New York: W. H. Freeman and Company. pp. 96-97); 덴드리머; 또는 양이온 중합체, 예컨대 DEAE-덱스트란 또는 폴리에틸렌아민을 포함한다. 비화학적 방법은 전기천공법; 초음파천공법; 및 광학 형질감염을 포함한다. 입자-기반 형질감염은 유전자 총의 자석 보조 형질감염을 포함한다(Bertram, J. (2006) *Current Pharmaceutical Biotechnology* 7, 277-28). 바이러스 방법이 또한 형질감염에 사용될 수 있다. mRNA 전달은 트랜스메신저(TransMessenger[®]) 및 트랜스아이티(TransIT[®])를 사용하는 방법을 포함한다(Bire 등 *BMC Biotechnology* 2013, 13:75). 이중 DNA를 세포 내로 도입하는데 통상적으로 사용되는 방법 중 하나는, 예컨대 Wigler 등에 의해 기술된 바(Proc. Natl. Acad. Sci. USA 77:3567, 1980)와 같은 칼슘 포스페이트 침전이다. 박테리아 원형질체와 포유류 세포의 폴리에틸렌-유도된 융합(Schaffner 등, (1980) *Proc. Natl. Acad. Sci. USA* 77:2163)은 이중 DNA를 도입하는 또 다른 유용한 방법이다. 전기천공법이 또한 DNA를 숙주 세포의 세포질 내로 직접 도입하기 위해 사용될 수 있는데, 이는 예컨대 Potter 등(Proc. Natl. Acad. Sci. USA 81:7161, 1988) 또는 Shigekawa 등(BioTechniques 6:742, 1988)에 기술된 바와 같다. 이중 DNA를 포유류 세포로 도입하는데 유용한 어떤 시약은 리포펙틴(Lipofectin[®]) 시약 및 리포펙타민(Lipofectamine[®]) 시약(Gibco BRL, Gaithersburg, Md.)을 포함한다. 이들 시약은 모두 지질-핵산 복합체(또는 리포좀)를 형성하는데 사용되는 상업적으로 이용가능한 시약으로, 배양된 세포에 적용시, 핵산을 상기 세포 내로 쉽게 흡수시킨다.

[0183] **항원 결합 단백질 제작 방법(Methods for Making Antigen-Binding Proteins)**

[0184] 본 개시는 이중 특이적 항원 결합 단백질의 제작 방법을 또한 제공한다. 본원에 개시된 방법을 사용하여, 원하는 항원 결합 단백질을 높은 역가 및/또는 높은 특이적 생산성(pg/세포/일)으로 생산할 수 있다. 일부 구현예에서, 항원 결합 단백질은 적어도 1 g/L, 1.5 g/L, 2.0 g/L, 2.5 g/L, 3.0 g/L, 3.5 g/L, 4.0 g/L, 4.5 g/L, 5.0 g/L, 또는 10 g/L 이상의 역가로 생산된다. 일부 구현예에서, 항원 결합 단백질은 일당 세포당 생산된 총 항원 결합 단백질(pg 단위)에 기초하여 결정된 적어도 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 피코그램/세포/일(pcd) 이상의 특정 생산성으로 생산된다. 일부 구현예에서, 항원 결합 단백질은 적어도 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 50% 또는 60% 이상의 이중 특이적 항원 결합 단백질 역가 대 총 항원 결합 단백질 역가의 비율로 생산된다.

[0185] 일부 구현예에서, 상기 방법은 본원에 개시된 시스템을 사용하고, 시스템에서의 벡터를 형질감염에 의해 시스템의 세포 내에 도입한다. 외인성 핵산이 2개의 발현 강화된 유전자좌 내에 RMCE를 통해 적절하게 통합된 형질감

염된 세포는 스크리닝되고 동정될 수 있다. 일부 구현예에서, 형질감염된 세포의 동정은 형질감염 전에 숙주 세포 내에 존재하는 하나 이상의 선별 표지에 대한 음성 선별에 의해 달성된다. 다른 구현예에서, 형질감염된 세포의 동정은, 통합되도록 설계된 벡터에서 핵산에 의해 제공된 하나 이상의 선별 표지에 기초한 양성 선별과, 형질감염 전에 숙주 세포 내에 존재하는 하나 이상의 선별 표지에 대한 음성 선별의 조합에 의해 달성된다. HCF 함유 폴리펩티드 및 LCF 함유 폴리펩티드는 통합된 핵산으로부터 발현될 수 있고, 항원 결합 관심 단백질은 동정된 형질감염 세포로부터 수득될 수 있고, 공지된 방법을 사용해 정제될 수 있다.

[0186] 또 다른 구현예에서, 상기 방법은, 항원 결합 단백질을 함께 암호화하고 세포로부터 상기 항원 결합 단백질을 발현시키는 2개의 발현 강화 유전자와 내에 통합된 외인성 핵산을 함유하는, 상기 본원에 기술된 세포를 단순히 사용한다. 각각의 클로닝된 발현 카세트(들)은 각각의 특이적 통합 부위 내에서 인접한다.

[0187] 본 명세서는 다음의 실시예들에 의해 추가로 예시되며, 이들 실시예는 어떤 방식으로든 제한하는 것으로 간주되지 않아야 한다. 모든 인용된 참조 문헌(문헌 참조, 발행된 특허 및 본 출원 전반에 걸쳐 인용된 공개된 특허 출원을 포함함)의 내용은 본원에 참조로서 명백하게 포함된다.

[0188] 실시예

[0189] 실시예 1: 2개의 특이적 발현 강화 유전자좌에서 (부위 특이적 통합을 통한) 단일 특이적 항체(Abs)의 발현

[0190] 도 1에 예시된 바와 같이, RSS 부위가 Ab 발현 카세트 및 선별 표지에 대한 발현 카세트의 측면에 위치하는 벡터 내에 Ab 사슬(AbC1, AbC2)을 클로닝하였다. 2개의 Ab 사슬을 별도의 벡터 내에 클로닝하거나, 2개의 발현 카세트가 가능한 순서 중 임의의 하나로 일렬 배열된 하나의 벡터 내에 결합할 수 있었다: AbC1, AbC2 및 선별 표지(예를 들어, AbC1은 통상의 LC와 동등하고 AbC2는 통상의 중쇄와 동등함).

[0191] 요약하자면, VH 및 VL 도메인을 암호화하는 DNA는 단일 항원 양성 B 세포로부터 PCR에 의해 직접 분리될 수 있다. 중쇄 및 경쇄 PCR 산물을 IgG 중쇄 불변 영역 및 카파 경쇄 불변 영역을 각각 함유하는 Sap I 선형화 항체 벡터 내에 클로닝하였다. 중쇄 플라스미드(AbC2)는 중쇄 발현 카세트의 측면에 위치하는 RRS3 및 RRS2 부위를 갖는다. 또한, 중쇄 플라스미드에서 RRS3의 바로 하류에는 분리된 선별 표지 유전자가 있다(예: US7582298). 중쇄 플라스미드는 경쇄 발현 카세트의 측면에 위치하는 RRS1 및 RRS3 부위를 갖는다. 또한, 경쇄 플라스미드는 RRS3에서 ATG의 직전에 강한 프로모터를 가지므로, 숙주 세포 유전자좌 내에 통합될 때, 경쇄 플라스미드로부터의 RRS3 근위 프로모터 및 개시 ATG가 적절한 관독 프레임에서 중쇄 플라스미드 내의 선별 표지 유전자에 인접하게 위치되어 선별 유전자의 전사 및 번역을 허용한다. 이어서, 중쇄 가변 영역 서열을 갖는 정제된 재조합 플라스미드 및 동일한 B 세포로부터의 경쇄 가변 영역 서열을 갖는 플라스미드를 조합하고, 서열번호 1(EESYR®; 1번 유전자좌) 및 서열번호 2 유전자좌에서 적절한 RSS 및 선별 표지를 갖는 변형된 CHO 숙주 세포주 내에, 재조합 효소를 발현하는 플라스미드와 함께 형질감염시켰다. 변형된 CHO 숙주 세포주는 2개의 전사적으로 활발한 유전자좌에서 4개의 상이한 선별 표지를 함유한다. 결과적으로, 선별 표지가 상이한 형광 표지인 경우, 생산 CHO 세포는 원하는 세포 재조합체를 나타내는 양성-음성 조합을 위한 유동 세포 계측법에 의해 분리될 수 있다. 중쇄 및 경쇄 유전자를 발현하는 재조합 플라스미드가 재조합 효소를 발현하는 플라스미드와 함께 형질감염되면, 재조합 효소에 의해 매개된 부위 특이적 재조합은 RRS 및 치환을 함유하는 각각의 염색체 유전자좌에서 항체 플라스미드를 통합시킨다. 따라서, 단일 특이적 항체를 발현하는 재조합 세포를 분리하고 12일 동안 유가 배양 제조한 뒤, 수확하고, 고정화 단백질 A를 사용하여 옥텟 역가 검정을 수행하였다. 세포는 동형 유전자이며 안정한 것으로 관찰되었다. 작은 셰이커 플라스크에서의 전체 역가는 2개의 부위 통합 방법을 사용할 때 단일 특이적 항체의 발현에 대해 증가하는 것으로 관찰되었으며, 항체 B의 경우 역가가 거의 두 배가 되는 상당한 증가를 나타냈다(도 2).

[0192] 실시예 2: 2개의 특이적 발현 강화 유전자좌에서 (부위 특이적 통합을 통한) 이중 특이적 항체(BsAbs)의 발현

[0193] 이중 특이적 항체의 발현을 위해, RRS 부위가 측면에 위치한 AbC1, AbC2 및 선별 표지 1이 제1 유전자좌 또는 통합 부위(EESYR®, 서열번호 1; 1번 유전자좌)와 양립할 수 있고, AbC1, AbC3 및 선별 표지 2가 제2 유전자좌 또는 통합 부위(서열번호 2)와 양립할 수 있도록 3개의 항체 사슬 및 2개의 선별 표지를 실시예 1과 유사한 플라스미드 내로 클로닝하였다. 관찰에 있어서, 통상적인 LC로서의 AbC1은 충분한 발현을 위해 2개의 유전자 카피를 필요로 하지 않는다. 각 부위마다 1개 또는 2개의 플라스미드를 만들었는데, 3개의 발현 카세트는 일렬로 배열되거나 2개의 플라스미드로 배열되며, 이중 2개의 발현 카세트는 하나의 벡터에 클로닝되고 나머지 발현 카세트가 제2 벡터에 클로닝된다. 도 3을 참조한다.

[0194] 중쇄 및 경쇄 유전자를 발현하는 재조합 플라스미드가 재조합 효소를 발현하는 플라스미드와 함께

형질감염되면, 재조합 효소에 의해 매개된 부위 특이적 재조합은 RRS 및 치환을 함유하는 각각의 염색체 유전자좌에서 항체 플라스미드를 통합시킨다. 따라서, 이중 특이적 항체를 발현하는 재조합 세포를 단리하고 12일 동안 유가 배양 제조한 뒤, 수확하고 고정화 항Fc 및 제2 항Fc*(변형된 Fc 검출 항체, 2014년 5월 15일 공개된 US 2014-0134719 A1 참조)를 사용해 옥텟 역가 검정을 수행하였다. 세포는 동형 유전자이며 안정한 것으로 관찰되었다. 작은 셰이커 플라스크에서의 전체 역가는 2개 부위 통합 방법을 사용으로부터 1.75배 내지 2배 초과로 상당히 증가한 것으로 관찰되었다(도 5).

[0195] 실시예 3: 부위 특이적 통합 후 이중 특이적 및 단일 특이적 항체의 대규모 생산

[0196] 전술한 실시예 1과 유사하게 숙주 세포(CHO-K1)를 생성하였다(이중 특이적 항체의 경우 도 3을, 단일 특이적 항체의 경우 도 1을 또한 참조한다). EESYR® 유전자좌(1번 유전자좌) 및 서열번호 2(2번 유전자좌)에서 유전자 카세트의 RMCE를 가능하게 하는 숙주 세포를 단 하나의 통합 부위(1번 유전자좌/EESYR®) 내로의 유전자 카세트의 RMCE를 가능하게 하는 숙주 세포와 비교하였다. 항체 경쇄 및 중쇄(AbC1, AbC2, AbC3)를 가지는 벡터, 필요 RRS 및 선별 표지 핵산(도 3 참조)을 생산 세포주(RSX^{2BP})로 재조합하여 Ab E, Ab F, Ab G, 및 Ab H를 발현하는 숙주 세포를 생성하였다. 이와 같이, 각각의 이중 특이적 항체 숙주 세포는 1개의 공통 경쇄 및 상이한 항체에 결합하는 2개의 중쇄를 발현하는데, 중쇄 중 하나는 그의 CH3 도메인에서 조작되어 단백질 A에 차별적으로 결합한다(본원에 참조로서 통합된 미국 특허 제8,586,713호에 기술되어 있음).

[0197] 단일 특이적 항체의 경우, 도 1에 예시된 바와 같이, RSS 부위가 Ab 발현 카세트 (및 선별 표지 유전자를 제공하는 발현 카세트)의 측면에 위치하는 벡터에 항체 경쇄 및 중쇄(AbC1, AbC2)를 클로닝하였다. 재조합을 수행하여 Ab J 및 Ab K를 발현하는 숙주 세포(RSX²)를 생성하였다.

[0198] CHO-K1로부터 유래된 항체 생산 세포주 RSX²의 시드 배양물로부터 2L, 15L, 또는 50L의 바이오리액터를 접종하였다. 접종된 세포를 36.5°C에서 13일 동안 성장시키고 필요에 따라 포도당과 기타 보충 영양분을 공급했다. 세포를 화학적으로 한정된 (가수분해물이 없고 혈청이 없는) 염기 배지에서 성장시켰다. 모든 항체를 수확하고 정제하였다.

[0199] 총 IgG 항체 (역가)를 단백질 A/단백질 G 크로마토그래피에 따라 측정하였다. 이중 특이적 항체의 경우, 원하는 이중 특이적 항체 종의 비율을 결정하기 위해 총 IgG 항체뿐만 아니라 이중 특이적 (이중 이량체 Fc/Fc*), 야생형 중쇄(Fc/Fc)를 갖는 동종 이량체 및 변형된 중쇄를 갖는 동종 이량체(Fc*/Fc)를 포함하는 3종의 항체를 측정하였다. 총 역가는, 미국 특허 제8,586,713호에 기술된 바와 같은 용리 기술을 사용하여 단백질 A/단백질 G 칼럼을 사용하는 HPLC 방법에 의해 결정하였다. 요약하자면, 3종의 바이오리액터는 시료가 로딩되는 동안 칼럼에 결합하고, 이중 특이적 종(Fc/Fc*)은 이온 조절자의 존재 하에서 pH 단계 구배를 사용하여 단백질 A 칼럼에서 먼저 용리된다. 이중 특이적 종이 제1 용리 단계에서 수집된 뒤, 2개의 동종 이량체 종이 용리된다.

[0200] 표 1은, 파일롯 대규모 생산 배양물에서 전반적 (총) IgG 역가 및 이중 특이적 항체 역가(도 6a)가 2개의 통합 부위를 통해 항체를 발현하는 숙주 세포를 사용함으로써 크게 개선되었음을 보여준다.

표 1

총 IgG 및 이중 특이적 IgG 역가 측정

항체	이중 특이적 역가 (g/L)		총 역가 (g/L)	
	2 개 부위 (1 번 유전자좌/2 번 유전자좌)	1 개 부위 (1 번 유전자좌)	2 개 부위 (1 번 유전자좌/2 번 유전자좌)	1 개 부위 (1 번 유전자좌)
Ab E	2.25	0.87	4.47	2.7
Ab F	1.45	0.7	2.9	2.4
Ab G	1.6	0.5	2.7	1.3
Ab H	2	1	3.2	2.5

[0201]

[0202] 전술한 바와 같이 이중 특이적 역가를 결정하였다. 표 2에 나타난 바와 같이, 세포에 의해 생산된 총 IgG 역가의 비율로서의 이중 특이적 항체의 역가는 2개의 통합 부위 구성을 갖는 숙주 세포의 생산 배양에서 상당히 더 높다. 도 6b 참조. 실제로, 50% 이상의 이중 특이적 비율이 일관되게 달성될 것으로 예상하지 않았다.

[0203]

표 2

총 IgG 생산당 이중 특이적 항체의 비율

	이중 특이적 비율 (% 총 IgG)	
	2 개 부위 (1 번 유전자좌/ 2 번 유전자좌)	1 개 부위 (1 번 유전자좌)
Ab E	50%	32%
Ab F	50%	29%
Ab G	59%	38%
Ab H	63%	40%

[0204]

[0205] 전반적 IgG 역가의 개선을 측정하는 2개 통합 부위 방법을 사용하여 단일 특이적 항체를 발현시켰다. 표 3은, 대형 바이오리액터 규모에서 전반적 역가가 0.6 내지 1.3배로 상당히 증가하는 것을 관찰하였음을 도시한다. 도 7을 또한 참조한다. 제조에 사용된 생산용 바이오리액터, 특히 배양 용적이 500 L에서 최대 10,000 L인 생산용 바이오리액터의 경우, 이들 개선된 세포주의 사용에 의한 역가의 증가는 배치 당 증가된 생산 수율의 상당한 양과 동일한 것으로 관찰되었다.

표 3

총 IgG (단일 특이적) 역가 측정

	총 역가 (g/L)	
	2 개 부위 (1 번 유전자좌/2 번 유전자좌)	1 개 부위 (1 번 유전자좌)
Ab J	3.3	2.2
Ab K	3.2	2

[0206]

[0207] 전술한 발명이 예시와 실시예로서 일부 상세히 설명되었지만, 당업자라면 첨부된 청구범위의 사상 및 범주를 벗어나지 않고도 본 발명의 교시에 특정 변경 및 변형을 가할 수 있다는 것을 쉽게 이해할 것이다.

[0208] 서열

[0209] 서열번호 1

[0210]	13515 염기	
[0211]	DNA	
[0212]	Cricetulus griseus	
[0213]	tctagaaaca aaaccaaataa tattaagtca ggcttggctt caggtgctgg ggtggagtgc	60
[0214]	tgacaaaaat acacaaattc ctggctttct aaggcttttt cggggattca ggtattgggt	120
[0215]	gatggtagaa taaaaatctg aaacataggt gatgtatctg ccatactgca tgggtgtgta	180
[0216]	tgtgtgtgta tgtgtgtctg tgtgtgtgcc cagacagaaa taccatgaag gaaaaaaca	240
[0217]	cttcaaagac aggagagaag agtgacctgg gaaggactcc ccaatgagat gagaactgag	300
[0218]	cacatgccag aggaggtag gactgaacca ttcaacacaa gtggtgaata gtcctgcaga	360
[0219]	cacagagagg gccagaagca ctcagaactc cagggggtca ggagtgggtc tctggaggct	420
[0220]	tctgcccttg gaggttctcg aggaggaggc ttccatattg aaaatgtagt tagtggccgt	480
[0221]	ttccattagt acagtgacta gagagagctg agggaccact ggactgaggc ctagatgctc	540
[0222]	agtcagatgg ccatgaaagc ctagacaagc acttccgggt ggaaaggaaa cagcaggtgt	600
[0223]	gaggggtcag gggcaagtta gtgggagagg tcttcagat gaagtagcag gaacggagac	660
[0224]	gcactggatg gccccacttg tcaaccagca aaagcttggg tcttgttcta agaggccagg	720
[0225]	gacatgacaa gggatgatctc ggtttttaaa aggctttgtg ttacctaate acttctatta	780
[0226]	gtcagatact ttgtaacaca aatgagtact tggcctgtat tttagaaact tctgggatcc	840
[0227]	tgaaaaaaca caatgacatt ctggctgcaa cacctggaga ctcccagcca ggccctggac	900
[0228]	ccgggtccat tcatgcaaat actcaggac agattcttca ctaggtactg atgagctgtc	960
[0229]	ttggatgcaa atgtggcctc ttcattttac tacaagtac catgagtcag gaggtgctgt	1020
[0230]	ttgcacagtg tgactaagtg atggagtgtt gactgcagcc attcccggcc ccagcttgtg	1080
[0231]	agagagatcc ttttaaatg aaagtaagct caaagttacc acgaagccac acatgtataa	1140
[0232]	actgtgtgaa taatctgtgc acatacacia accatgtgaa taatctgtgt acatgtataa	1200
[0233]	actgtgtgaa taatctgtgt gcagcctttc cttacctact accttccagt gatcaggttt	1260
[0234]	ggactgcctg tgtgtactg gacctgaat gtccccaccg ctgtcccctg tcttttacga	1320
[0235]	ttctgacatt ttttaataaa tcagcggctt cccctctgct ctgtgcctag ctataccttg	1380
[0236]	gtactctgca ttttggtttc tgtgacattt ctctgtgact ctgctacatt ctcagatgac	1440
[0237]	atgtgacaca gaaggtgttc cctctggaga catgtgatgt cctgtcatt agtggaatca	1500
[0238]	gatgccccca aactgttgtc cagtgtttgg gaaagtgaca cgtgaaggag gatcaggaaa	1560
[0239]	agaggggtgg aaatcaagat gtgtctgagt atctcatgtc cctgagtggt ccaggctgct	1620
[0240]	gacttcactc cccaagtga gggaggccat ggtgagtaca cacacctcac acatactata	1680
[0241]	tccaacacac acacacacac acacacacac acgcacgcac gcacgcacgc acgcacacat	1740
[0242]	gcacacacac gaactacatt tcacaaacca catacgcata ttacaccca aacgtatcac	1800
[0243]	ctatacatat cacacatata caccctcca cacatcacac acataccaca cccacacaca	1860
[0244]	gcacacacat acataggcac acattcacac accacacata tacatttgtg tatgcataca	1920
[0245]	tgcatacaca cacaggcaca cagacaccac acacatgcat tgtgtacgca cacatgcata	1980

[0246]	cacacacata ggcacacatt gagcacacac atacatttgt gtacgcacac tacatagaca	2040
[0247]	tatatgcatt tgtatatga cacatgcatg cacacataca taggcacaca tagagcacac	2100
[0248]	acatacatatt gtgtatgcac acatgcacac accaatcaca tgggaagact caggttcttc	2160
[0249]	actaaggttc acatgaactt agcagttcct ggttatctcg tgaacttgg aagattgctg	2220
[0250]	tggagaagag gaagcgttgg cttgagccct ggcagcaatt aacccgccc agaagaagta	2280
[0251]	ggtttaaaaa tgagagggtc tcaatgtgga acccgaggg cgccagttca gagaagagac	2340
[0252]	ctaccaagc caactgagag caaaggcaga gggatgaacc tgggatgtag ttgaaacctc	2400
[0253]	tgtaccagct gggcttcatt ctattttgtt atatctttat taaatattct tttagtttta	2460
[0254]	tgtgcgtgaa taccttgcct gcataaatgt atgggcactg tatgtgttct tgggtccggt	2520
[0255]	ggaggccagg agagggcatt gatcctccgg agctggcgtt tgagacagtt gtgaccaca	2580
[0256]	gtgtggggtc tgggaactgg gtcttagtgt tccgaagtg cagctggggc tcttaacctc	2640
[0257]	tgagccatcc ctccagcttc aagaaactta ttttcttagg acatggggga agggatccag	2700
[0258]	ggctttaggc ttgtttgttc agcaaatact ctttcttgtt atttgaatt ttattttatt	2760
[0259]	ttactttttt gggatagaat cacattctgc agctcaggct gggcctgaac tcatcaaat	2820
[0260]	cctcctgtct cagctacca ggtgataaga ttactgatgt gagcctggct ttgacaagca	2880
[0261]	cttttagagtc ccagccctt ctggacactt gttccaagta taatatatat atatatatat	2940
[0262]	atatatatat atatatatat atatatgtg tgtgtgtgtt tgtgtgtgta tgagacactt	3000
[0263]	gctctaaggg tatcatatat atccttgatt tgcctttaat ttatttttta attaaaaatg	3060
[0264]	attagctaca tgcacctgt atgcgtctgt atcatctata tctcttcct tccttctctc	3120
[0265]	tctttctctc tcttctctc ccccccaag catctatitt caaatcctt tgccaggag	3180
[0266]	atgccaagag tctcgttggg ggagatggtg agggggcgat acaggggaag agcaggagga	3240
[0267]	aagggggaca gactggtgtg ggtctttgga gagctcagga gaatagcagc gatcttcct	3300
[0268]	gtccctggtg tcacctctta cagecaacac catittgtgg cctggcagaa gattgtcaa	3360
[0269]	gctggtcgca ggctgccac acaaccccaa tctggcccca agaaaaggca cctgtgtgtg	3420
[0270]	actctggggt taaaggcgt gcctggctgt ctccagctgg acttgaaact cccgtttaat	3480
[0271]	aaagagtctt gcaaaataat acccgagag tcacagtgc aggttccgt gctttcctga	3540
[0272]	agcgccaggc acgggttccc taggaaatgg ggccttgctt gccagctcc cacggcttgc	3600
[0273]	cctgcaaacg gcctgaatga tctggcactc tgcgttgcca ctgggatgaa atggaaaaaa	3660
[0274]	gaaaaagaag aagtgtctct ggaagcgggc gcgctcacac aaacccgcaa cgatttgtta	3720
[0275]	aacactctcc attgagaatc tggagtgcgg ttgccctcta ctggggagct gaagacagct	3780
[0276]	agtggggcg gggggaggac cgtgctagca tccttccacg gtgctcgtg gctgtggtgc	3840
[0277]	atgccgggaa ccgaaacgag gaactaaagt caagtcttgc tttggtggaa ctgacaatca	3900
[0278]	acgaaatcac ttcgattgtt ttctcttttt tactggaatt cttggatttg atagatgggg	3960
[0279]	gaggatcaga gggggagggg aggggcgggg agacggaggg aggaggggag gaggggagga	4020
[0280]	ggggaggagg ggaggagggg aagggatgga ggaataact aacttttcta attcaacatg	4080
[0281]	acaaagattc ggagaaagtg caccgctagt gaccgggagg aggaatgccc tattgggcat	4140

[0282]	tatattccct gtcgtctaata ggaatcaaac tcttgggtcc agcaccaagg attctgagcc	4200
[0283]	tatcctattc aagacagtaa ctacagccca cacggaagag gctatacaac tgaagaaata	4260
[0284]	aaattttcac tttatttcat ttctgtgact gcatgttcac atgtagagag ccacctgtgt	4320
[0285]	ctaggggctg atgtgctggg cagtagagtt ctgagcccgt taactggaac aaccagaac	4380
[0286]	tcccaccaca gttagagctt gctgagagag ggaggccctt ggtgagattt ctttgtgtat	4440
[0287]	ttatttagag acagggtctc atactgtagt ccaagctagc ctccagctca cagaaattct	4500
[0288]	cctgttccgg tticcaaagt actggagtta tgagtgtgtg ttaattgaac gctaagaatt	4560
[0289]	tgctgattga agaaaacctc aagtgggttt ggctaataccc cagacccca gaggtgagg	4620
[0290]	caggaggaat gagagaattc aaggtttggc agagccacag ggtgagctca atgtggagac	4680
[0291]	tgtgagggtg agctcaatgt ggagactgtg agggtgagct caatgtggag actgtgagg	4740
[0292]	tgagctcaat gtggagactg tgagggtgag ctcaatgtgg agactgtgag ggtgagctca	4800
[0293]	atgtggagac ctgtatcaag ataataatag tagtagtaac aatgcaggcg aggggtgtgt	4860
[0294]	tgagtggtag agcagttagt tgatttgaca tgcttgaggt ctcccgttc atctgtggcc	4920
[0295]	ctgcaacagg aaggaggga ggaagggggg gaacgagaga gagaaagag agacagaagc	4980
[0296]	taagataggg aatgagagag gaaggaagaa acgggaagaa attcagactc cttcctgagt	5040
[0297]	tccgccaacg cctagtgaca tcctgtgcac accctaaggt ggccctttgtg tggcactggc	5100
[0298]	ttgggtggtc gggaaaggca ttttcagctt gttgcagaac tgccacagta gcatgtggg	5160
[0299]	tccgtgaaag tttctgccc ttaacaagaa gtcttacta cttgtgacct caccagtga	5220
[0300]	aatttcttta attgtctcct ggtgttctgg gttttgcatt tttgtttcta aggatacatt	5280
[0301]	cctgggtgat gtcatgaagt ccccaaagac acagtggggc tgtgttggat tgggaaagat	5340
[0302]	gatttatctg ggggtgtcaa aggaaaagaa gggaaacagg cacttgggaa aatgtcctcc	5400
[0303]	cgcccaccg aattttggct tggcaaccgt ggtggaggag caagaaacac gtggacgttt	5460
[0304]	gaggaggcat ggggtcctag gaggacagga agcagaagga gagagctggg ctgacagcct	5520
[0305]	gcaggcattg cacagtttca gaaggagatt acagcatgac tgagttttta gggatccaac	5580
[0306]	agggacctgg gtagagattc tgtgggctct gaggcaactt gacctcagcc agatgttatt	5640
[0307]	tgaataacct gctcttagag ggaaaacaga catagcaaac agagccacgt ttagtgatga	5700
[0308]	aactctcact ttgcctgagt catgtgcggc catgccagg ggtcaggctg aactcaact	5760
[0309]	caaaaacaag tgagaaattg aagacaatcc gtggtggcag ctactggaag ggccaccaca	5820
[0310]	tcccagaaa gagtggagct gctaaaaagc catttgtgat aggcacagtt atcttgaatg	5880
[0311]	catggagcag agattacgga aaaatcgaga atgttaatga ggcaacattc gagttgagtc	5940
[0312]	attcagtgtg gaaaccag acgcttccat cccctaaaag gaacatcttg ctctcagtc	6000
[0313]	aatggaaat aaaaattggg gcttgaattt ggcaaatgat tcagaactct gtgtaggtat	6060
[0314]	tttcacagc acagtggata attttcatgt tggagtttat ttgtgctaaa aggcagaaaa	6120
[0315]	gggtaaaaag cacatcttaa gagttatgag gttctacgaa taaaaataat gttacttaca	6180
[0316]	gtatttcctt aattagtacc cccttcacc tgtggtaatt tcctgagata gtcagtgggg	6240
[0317]	aaaagatctc tccttctctt ctttctcccc ctccctcct ctccctcct cctccctcc	6300

[0318]	ctccctcctc tccctccctc cccctttcct tctttctttg ctccctctcc tctgcctcct	6360
[0319]	tctccctttc ttcttcattt attctaagta gcttttaaca gcacaccaat tacctgtgta	6420
[0320]	taacgggaaa acacaggctc aagcagctta gagaagattg atctgtgttc actagcgtgc	6480
[0321]	aattcagagg tgggtgaaga taaaaggcaa acatttgagg ccatttcctt atttggcacg	6540
[0322]	gcacttagga agtggaaacat gcctaattca ctggtttgta ccacctttcc ctataatgga	6600
[0323]	ctgtttggga agctcctggg caaccgattc tggcatctca ttggtcagag gcctgttaaa	6660
[0324]	tgggtactctt atttgcaaag aaggctgtaa ctgttagctt taaaagcctc tcctcaagaa	6720
[0325]	agaagggaga aaggataatg ctagacatat ctaatagact taaccactgt gaaaagcctt	6780
[0326]	agtatgaatc agatagaacc tatttttaac tcagttttga aaaaaataat ctttatattt	6840
[0327]	atttgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gaaccacatg	6900
[0328]	tagcagggtgc tggaggaggc cagaagaggg caccagatct cctggaactg acaccacaca	6960
[0329]	tggttatgag ctgcctgatg tgggtgctgg gaactgaact ctctgtttct gcaagagcag	7020
[0330]	caactgttct cttaactgat gagccatctc tccagccccc ccataattt taattgttca	7080
[0331]	tttttagtaa ttttattcat aatcaattat cacagtataa aacaatgatt ttatatatat	7140
[0332]	catatacata tcaaggatga cagtgagggg gatatgtgtg tgtgtgtgtg tgtgtgtgtg	7200
[0333]	tgtgtgtgtg tgtgttattt gtgtgtgtgc tttttaagaa ggtgccatag tctactgcatt	7260
[0334]	tctctgaagg atttcaaagg aatgagacat gtctgtctgc caggaaacct atcttctct	7320
[0335]	ttgggaatct gacccaaatg aggtattctg aggaactgaa tgaagagctc aagtagcagt	7380
[0336]	gtcttaaacc caaatgtgct gtctagagaa agtcaacgtc atcagtgagc tgaggagaga	7440
[0337]	tttactgagc ggaagacaag cgctctttga ttttaagtgc tcgaacagtc acggctgtgg	7500
[0338]	agtggagcct gtgctcaggt ctgaggcagt ctttgctagc cagctgtgat gagcagtga	7560
[0339]	gaaagggtgg agatggaggc aggggtgggag cagggtctatg gttcagacta ggtatcgtga	7620
[0340]	gcacaccagc tggttgactt gtggctctgt ggtcaggcgt tgtaaacgcc ctgagggtca	7680
[0341]	ggcagtcaca ttgcttgaag ctgaatgggt gaggcaacac agagagtgc aagaaggcaa	7740
[0342]	agtaccacct cttccccgac ccaggtcact tctgggttat agctgagact ccggacagca	7800
[0343]	tgcaaccagc tggttagagc ttcagggaaa acttgatgtc tgcatgttgc tatgaaatgt	7860
[0344]	gattcggtag atctggagaa aatttataat gctggctcag tcaagcactg aacaaaggta	7920
[0345]	ccttggcttt gggagctaca tgacattgac ttgtaggcag actttttttt ttctgcccgc	7980
[0346]	caattccagc ataaccaata tggaggctca atattaatta taaatgctcg gctgatagct	8040
[0347]	caggcttggt actagctaac tcttccaact taaatgaacc catttctatt atctacattc	8100
[0348]	tgccacgtga ctttaccttg tacttctgtt ttctctcct tgtctgactc tgcccttctg	8160
[0349]	cttcccagag tecttagtct ggttctctg cctaacctta tctgcccag ctgctgacca	8220
[0350]	agcatttata attaatatta agtctcccag tgagactctc atccaggag gacttgggtg	8280
[0351]	ctccccctc ctcatgtcca tccgtgtctt cctcttccct cgttcccc tctcttct	8340
[0352]	gtcttctcct ctccaccct ctttcatag tattgatggc aagggtgttc tagaatggag	8400
[0353]	gagtgccc ataggcatgcaa agaaaccagt taggatgctc tgtgagggtg tgtaatcata	8460

[0354]	agcgatggac acaattcaag ccacagagtg aagacggaag gatgcactgt gctctagagc	8520
[0355]	aacttctggg gcagaatcac aggggtgagtt tctgacttga gggcgaagag gccacgagga	8580
[0356]	agggagttag tttgtctgag ctagaagcta cgccccacct cttggttagca gacctgcccc	8640
[0357]	caagcatgct ttgttaatca tgtgggatct gattttcttc taaatctatg ttcaactctt	8700
[0358]	aagaaaatgt gaattctcac attaaaattt agatatacgt cttttggtgg ggggggtgta	8760
[0359]	aaaaatcctc aagaatatgg atttctgggg gccggagaga tggctcagag gttaaagaaa	8820
[0360]	ctggttgctc ttctagacat tctgagttca attcccagca accacatggt ggctcacaac	8880
[0361]	catctgtaat gcgacctggt gccatcttct gacatgcatg gatacatgca ggcagaaagc	8940
[0362]	tgtatacata gtaaattgat aaatcttttt ttaaaaagag tatggattct gccgggtgtt	9000
[0363]	ggtggcgcac gcctttaatc ccagcactct ggaggcagag gcaggtggat ctctgtgagt	9060
[0364]	tcgagaccag cctgggtctat aagagctagt tccaggacag cctccaaagc cacagagaaa	9120
[0365]	ccctgtctcg aaaaaccaa aaaaaaaaaa aaaaaaaaaa aaaaaaaga gtatggattc	9180
[0366]	taagaaagcc gtaacagctg gagctgtgta cggagttcag cgtgggtaact gaagaacaga	9240
[0367]	cattcatgat gaaacacccc aggattttta cttagtatct agtttccatt gttgttttga	9300
[0368]	gaccggctct tatgctctcc aggcctggcct caaactgctg atcttcccgc ctctacctct	9360
[0369]	caagtcttgg gactacttgg ctcataaaac agtttttgtc gggctccctg aagttatggt	9420
[0370]	tgtacaaacc gtgggggtca atatactcac ttgggcagag agagaaggtc tgaatcccag	9480
[0371]	acaatgactg catctcagga cagtttggga gaggacaatg gcagaaggac ttagaaaaga	9540
[0372]	tagactggag ggtggaaaag cagcaggaac agagaaacaa aacaggaagc ttgctatcca	9600
[0373]	gggccactct ggagtctgtt ggcaagatgg aagcgggcta ggggaataca tttgtgtctac	9660
[0374]	tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgat caatgcctat caatgttgaa	9720
[0375]	ggggaaatat gtataccaca ttgattctgg gagcaattct cagtatctgg cctagagaaa	9780
[0376]	ggaatggccc ctgcagaata gacagagtga atgggtgccct ttatcatttg ctaaaagtga	9840
[0377]	ggagaaataa acatccttcc atagagtttc aggtaaatga accccacagt tcactgtgct	9900
[0378]	cgtggtggag gcctggccaa cagttaaaaa gattagacac ggacaaagtc tgaaggaaac	9960
[0379]	acctcgaata ggaagaggag agccacctca ttctgtaact ttctcaagg ggaagatgtt	10020
[0380]	ccaagagtgg gaataaatgg tcaaaggggg gatttttaat taggaaaacg atttctgtga	10080
[0381]	tcacttgtga aactggaggt tgatttgggg cataggacaa tagatttgat gctttgcaaa	10140
[0382]	aagctgtttc aaagcagaga aatggaatag agacaattat gtagcgagga gggagggtgg	10200
[0383]	ggcgaagatg gagacagaga agtgggaagct gactttaggg aagaggaaca tagaccacag	10260
[0384]	gggcggggcg gggggcaggg gcggggggcg gggctcaaag gaggcagtgg gaacgttgct	10320
[0385]	agtgttcgca gcgtaagcgt gaatgtgcaa gcgtctttgt ggtgtgtgac caggagttagc	10380
[0386]	gtggctggct tgtgtgctgc ttgtaatccc agtccttgag gtttccacac tgttccacag	10440
[0387]	tgggtgtgat tttccctcgg agagcatgag ggtctctgct tccccacatc ctccccagcg	10500
[0388]	ttcgttggta tttgtttcca agatgttagt gggtagagaa aagcctctct gttgatttgc	10560
[0389]	ctttaacagg tgacaaaaaa agctcaacca ggagacattt ttgccttctt ggaaggtaat	10620

[0390] gctcccatgt agagcaatgg gacccatctc taaggtgagg ctactcttgc agtttgcacc 10680

[0391] cagctcttct gatgcaggaa ggaagtgtgt gggcaagcaa gactgtttgc tttctgcgat 10740

[0392] ggacacattc tgcacacaaa ggctcaggag gggagaaggc tgtttgatgt ttagcactca 10800

[0393] ggaaggcccc tgatgcatct gtgattagct gtctccatct gtggagcaga cacggactaa 10860

[0394] ctaaaaacca gtgtttttaa attgtcaagc ctttaagggt aggaaattga cttattgtgc 10920

[0395] tgggccatac gtagagcaag tgctctgcat tgggccaacc cccggctctg gtttctaggc 10980

[0396] accagaatgg cctagaacta actcacaatc ctccattcc aggtctcagg tgctagaatg 11040

[0397] aaccactata ccagcctgcc tgctcgcta cctgccttcc taaattttaa atcatgggga 11100

[0398] gttagggaga atacacttat cttagttagg gtttctattg ctgtgaagag acaccatgag 11160

[0399] catggcaact cttataaagg aaaacattta gttgggtggc agtttcagag gttttagtac 11220

[0400] attgtcatca tggttgggaa catgatggca tgcagacaga catggtgctg gagaaaggga 11280

[0401] tgagagtcc t acatcttga ggcaacagga cctcagctga gacactggct ggtaccctga 11340

[0402] gcataggaaa cctcacagcc caccctcaca gtgacatatt tccttcaaca aagccatacc 11400

[0403] tcctaatagt gccactccct atgagatgac agggccaatt acattcaaac tgctataaca 11460

[0404] ctttaaagta ttttattttt attattgtaa attatgtatg tagctgggtg gtggcagccg 11520

[0405] aggtgcacgc ctttaatccc agcacttggg aggcagaggc agatggatct ctgtgagttc 11580

[0406] aagaccagcc tggctctataa gagctagtgt caaggaagga tatacaaaga acagtcttag 11640

[0407] gatagccttc aaagccacag agaagtgtgt tcttgaaaac caaaaattgt gctgggacct 11700

[0408] gtctctgctt tggttgcttc ccactcccc agagctggac tcttgggtcaa cactgaatca 11760

[0409] gctgcaaaat aaactcctgg attcctctct tgtaacagga gcccgaagtc aggcgcccac 11820

[0410] ttgtcttctc gcaggattgc catagacttt tttctgtgtc ccaccattcc agactgaagt 11880

[0411] agagatggca gtggcagaga ctgggaaggc tgcaacgaaa acaggaagtt attgcaccct 11940

[0412] gggaatagtc tggaaatgaa gcttcaaaac ttgcttcatg ttcagttgta cacagactca 12000

[0413] ctcccagggt gactcacacg tgtaaatatt cctgactatg tctgcaactgc ttttatctga 12060

[0414] tgcttccttc ccaaaatgcc aagtgtacaa ggtgagggaa tcacccttgg attcagagcc 12120

[0415] cagggtgctc ctcttaacc tggacttgtc tttctcggc agcctctgac accctcccc 12180

[0416] ccattttctc taticagaagg tctgagcaga gttggggcac gctcatgtcc tgataactc 12240

[0417] cttgtcttcc tgaagatcta acttctgacc cagaaagatg gctaaggtgg tgaagtgttt 12300

[0418] gacatgaaga cttggtctta agaactggag caggggaaaa aagtcggatg tggcagcatg 12360

[0419] taccgaaat ccagaactg gggaggtaga gacggatgag tgcccggggc tagctggctg 12420

[0420] ctgagccagc ctgctgaat tgccaaattc caactcctat tgaaaaacct ttaccaaaaa 12480

[0421] aacaaacaaa caaataataa caacaacaac aacaacaaac taccctatc aaggtgggag 12540

[0422] gctcttggct cttgaggaat gactcaccca aacccaaagc ttgccacagc tgttctctgg 12600

[0423] cctaaatggg gtgggggtgg ggcagagaca gagacagaga gagacatgac ttcctgggct 12660

[0424] gggctgtgtg ctctaggcca ccaggaactt tctgtcttg ctctctgtct ggcacagcca 12720

[0425] gagcaccagc acccagcagg tgcacacacc tcctccgtg cttcttgagc aaacacaggt 12780

[0426]	gccttggctt gtctattgaa ccggagtaag ttcttgcaga tgtatgcatg gaaacaacat	12840
[0427]	tgtcctgggtt ttattttctac tgttgtgata aaaaccgggg aactccagga agcagctgag	12900
[0428]	gcagaggcaa atgcaaggaa tgctgcctcc tagcttgctc cccatggctt gccgggcctg	12960
[0429]	ctttctgcaa gccctttctc ccccatgggc atgcctgaca tgaacagcgt ttgaaatgct	13020
[0430]	ctcaaatgct actttcaaag aaggctttct tgatcttgct aactaaatca gaccatgttt	13080
[0431]	caccgtgcat tatctttctg ctgtctgtct gtctgtctgt ctgtctatct gtctatcctc	13140
[0432]	tatcaatcat ctatctatct atcttctatt tatctaccta tcattcaatc atctatcttc	13200
[0433]	taactagtta tcattttatt atttgtttac ttactttttt tatttgagac agtattttctc	13260
[0434]	tgagtgcag ccttggtgt cctggaaccc attctgtaac caggtgttcc tcaaactcac	13320
[0435]	agagatccaa ctgcctctgc ctctctgggt ctgggggtta agacgtgcac caccaacgcc	13380
[0436]	ccgtctatc atctatttat gtacttatta ttcagtcatt atctatctc taactatcca	13440
[0437]	tcctctgtct atccatcctc tatctatcta tctatctatc tatctatcta tctatcctc	13500
[0438]	atctataatc aattg	13515
[0439]	서열번호 2	
[0440]	14931 염기	
[0441]	DNA	
[0442]	<i>Cricetulus griseus</i>	
[0443]	기타_특징 (2176)..(2239)	
[0444]	n은 a, c, g, t이거나 뉴클레오티드가 없음	
[0445]	catgtacact tatgcaagta tgatatggcc caacacagta ttttacacca atttttatct	60
[0446]	ataaaatata catgtacatc aaaatatatt attaataata acatcattat tctttctttc	120
[0447]	caagtaataa acacatacac tgaaattttg gttcttgttg ataattttaa tgaaacagga	180
[0448]	aatgcaaatt tatcttagca tgtttacttc actttctttg catagataac cagtaatcac	240
[0449]	attgatggat catgtagtga aatgtatttt taggtatcta aggaattttg gcttcgtttt	300
[0450]	gtgcttgttg aactgaatt ctattcctaa caacagtgtg taaggattct gtctgatttc	360
[0451]	ttttaccagt atttgtccat ttgcattttc tttattatc atggtctgtg ttctagaaag	420
[0452]	tggaaggtag tgtgtcaagt ctgtttaaca tgtttccctg atgatcagtg tcttaacacc	480
[0453]	tctctgagta catgttggcc aatgtcgttt ctagacccat ctattcttgc ttgacttate	540
[0454]	ctggtacatg cctgccaaga aatttctcct cctcctttct gtctcttcac tgatttactt	600
[0455]	gatgtgtgga ttacacattg atcatatgga aatagaagat acaattttct ttattcacag	660
[0456]	tttggaagac ttcaatctc atagatcctc attatttttt gctactgttc cctatgctat	720
[0457]	ggtgaaattt ccatittgaat aattgcttaa acaattaaca agaaagaatc tatttttact	780
[0458]	tgcaataact tccatttcag aacatttact acactgttac tataatccaa aactagtatt	840
[0459]	atataatcat tgagaaatga ctaattcata atttggccat gacatttttt tcagaaacag	900
[0460]	aaaaagtgc caatacatc acaatgctat aaatattaag acttcagcaa attaaatatt	960
[0461]	tattcatgat atcacataaa attcatttat tatgttttat ttaaatgtgt ttttaaaca	1020

[0462]	gtggatcac taaatattaa gttagatgtg tttatgtgct taatgaattt atattttaga	1080
[0463]	atgttataag ttgtatataag tcaaatatgt aataaatttt attttttagg tctttctcat	1140
[0464]	taaggatattt taatttttggg tcccttttcc agagtgactc tagctcatga tgagttgaca	1200
[0465]	taaaaactaa acagtacaaa atgtacattg cattcagtat tgcacttgat ctttgactg	1260
[0466]	aagtttgagt cagttcatatc atttagtact tgggaagtac attaagctaa ctttcattgc	1320
[0467]	tctggcaaaa tgctcgataa gataagagtc tatgttgga agccatggca gcaggaaagt	1380
[0468]	aagactgctg atgatgttta atccatagtc aagacgcaga aggagatgaa tgctggatc	1440
[0469]	caacattttt tgctgttcat tttctctaga accctagtcc ataaagatgt atgacttgca	1500
[0470]	ttcaaaatgc gtccccctca gttgttcaac ttttctgtaa atatcctttc aggcatgtct	1560
[0471]	agaagattgt ttcgcaataa cttctcaatc cattcaagtt gatagtgcag attaatcact	1620
[0472]	gcagaataaa agcctgtaac ttggctcacg tgccaaggaa tatgcacact cctgacacat	1680
[0473]	caataagtaa atcaaagtgt agcttttggc tttaacattg ccagacttat gtaatgttct	1740
[0474]	gcacgttctt cctccatcac tttttattct aatgggtgtt ccttgacatt gaatcacgct	1800
[0475]	gtggaagctg cttagaatta acattgaaat ctactgatat atttatgatg cagcaattta	1860
[0476]	gatttactat tttacttaga attttttata attgagagaa tataatattt tcacagttat	1920
[0477]	ctatctgctg taaatagagg attttaaaaa aaatctctat aacttttttt tacaacacac	1980
[0478]	agtaaaatta agttaaaatt taataaagtc actatgttga tttcaaagtg tgctacgccc	2040
[0479]	acggtggta cgcagggtga gcagaagatg ccactaaggt gggctaaggc cgatgggttg	2100
[0480]	gggtctgcgc tccctggaga tgagccccag gcgggttcct ggcaatcagc tgcgatcatg	2160
[0481]	atgccgatg agccannnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	2220
[0482]	nnnnnnnnnn nnnnnnnnnc tgggtgactt tatggaaaga atttgataga tttcatgatg	2280
[0483]	tagaagaatt ttattaggct tattttacag gagactaaga ccctgggacc taaagatc	2340
[0484]	tgggtcctga gaatcaggaa atgggttagag acgtgggtga tggatgaga cagatttttag	2400
[0485]	agaactctta gatcatgggc aatgaccgca atctgatgct tagaatagat catctataaa	2460
[0486]	caattatgct gttctttttc tttctgttgt atgatctgat gatgtagccc ccttgccaag	2520
[0487]	tccctgata ccccttgcca agttccctga ttgtaacagt atataagcat tgcttgagag	2580
[0488]	catattcaac tacattgagt gtgtctgtct gtcatttcct cgccgattcc tgatttctcc	2640
[0489]	ttgagccttt tcccttgctc tccctcggtc ggtgggtctcc acgagaggcg gtccgtggca	2700
[0490]	aaagtgtata aatgttctaa aacatttgaa ctctaaaaca tgcaaaatga aaaattaaaa	2760
[0491]	taaataaaca tgaaaattaa aatatattag ctgctaaaag ttaacaata ctatataata	2820
[0492]	ttttgttatt agaattcaaa atcacattag ttggatttaa tttgaacatt gcattctttc	2880
[0493]	aataataatt tcaataaaaa aagtttcccc atgatatag aaaataataa catatgtatc	2940
[0494]	tatctattta tttaactaca catatatagc atttgtttca actaaaataa atgaatgagc	3000
[0495]	aaagcaccta agtaattggt gtctattata tttatgaagc caatagtttc aaataaatta	3060
[0496]	tcatgcataa ggaggatttg caaatgttaa accttttttg aaacagatat tcccagttac	3120
[0497]	agaaattata atttctaate tttctataa gtagaatgat gataattaat ataggccatt	3180

[0498] tgtaaataat gttcagatta aaatatcttc tatttcacta gagaagaatg atattaaatg 3240

[0499] tattatattt tatttcccat ttgtttgca ccactattct atatccctca gcagtttaaa 3300

[0500] ttgtttcac catatgtgtg tgtgtttgta tcttaaatat ggcaactaaa ttagaataat 3360

[0501] ttaataataa tctttaggag aaaagatatt gaattatttt atgttgatag gaaaatatct 3420

[0502] ttttaattgc caagaatact ttttcttcta ttttaggact gatcagaccc aggactaata 3480

[0503] ttttatatgt actaattcta tgtacaaaa tatgttatta tctcatgaat tctgtctcaa 3540

[0504] tattgaggta ataaaaatag tccatcatga actttaaaat taaaataatg attaatat 3600

[0505] ttttattcat atttgtttg tatgaatgg tatacatcac atgtgtgcct ggtgactgtg 3660

[0506] aatgtcagga gaaggtatga aagccactgg aattggaata agagataata ttigagatgt 3720

[0507] tatgtgggtg ctgagaatta gacgaagcc atcttcaaga atagccagca tactatacca 3780

[0508] ctgagtaatc cattcatccc tcaataatta tctttgtaga cagtaaatat atttctaaac 3840

[0509] tataaatgac cagaaaaatt aatgtattat taatgaagac attcatctca tgtgacacac 3900

[0510] ttcacctgtc taaatcagta acactctctc cactaattaa gattttctaa gtgcatgaca 3960

[0511] cttactattt ctaaagctgt ccaatggggg ccagtcacca gtcagaccc agtgagataa 4020

[0512] tccatgaatg catttatatc ttaggaaaaa ttcttatcta tgtagtattt agaacatttt 4080

[0513] catgtgaggg gataaacaag gaagcacaga tgctttctga tagaaacttt ctctttaatt 4140

[0514] catctagaaa aaaaaaacct ctcaggaaaa tctctcttgc tctcctccca atgctctatt 4200

[0515] cagcatcttc tcctactta attctagatc tttttctcta tgccctcttg ctgctgccct 4260

[0516] gctggctctg ctctatgcct ccccatgtca cttttctttg ctatctcacc gttaccttct 4320

[0517] ctgccctact ctctgccctc ttctctgctt ctcacatggc caggctctgg acaattatag 4380

[0518] ttatatgtta cattctcata acacatgata tgtcacatag tttctctcag gctagggata 4440

[0519] tcacaatgac tggccaatga gcaagtggcc ttgcatgtag ctctaagttg gtgatggttc 4500

[0520] ccagacagta agtagccatt tggttgaaat ttgaggttgg gtagtacatg aagactgaat 4560

[0521] tttcttcaa ctctggcctt gaaatagtaa aacaacacct atgaaaatga cgacctgtat 4620

[0522] ttgtctttag aggcaaccac atattgtctg cagggcctgc tttgaatttg ctctgaagtt 4680

[0523] agcttgtttg tgtaaaagga agaatcctat atcagcctga gaaatgtaaa atatcctagc 4740

[0524] atttcaagtc atcaaaatta tatggagagt ataaatcatc cttctgacta ttcatagtca 4800

[0525] tatttgtgtc caccaagtat aaaacacact accaaagggc tgtggaaaaa atcgccataa 4860

[0526] ctgttcttat tagggaggca tagcagtgg tctgaggaa gttacagcaa caaccagtca 4920

[0527] tccagtcaat aaccccatgg ctttgccact tggaggtacc caataatgtt tggctttgcc 4980

[0528] gagtaggact ccaacaaatt cagaggttca atttttaaat gctggttgtc actgctgaac 5040

[0529] agtcccatg cctctgcat aattccacaa tggaaagctt ttacactga ttgccaatca 5100

[0530] ttaaacagcc tactcagcat aaacaggtat gatattattc tgcattttgt tacattacta 5160

[0531] gatgaattcc tatttcttcc tacaatagtg gaactgaaaa aagatacaca atcactactac 5220

[0532] ccctctacta atcttatgac ttatatcatt tcaattttca gaccataatg caaactattg 5280

[0533] accaaaacat gtgaagatga aaaatagaaa tgtagaataa tattacatat aaaaagaaaa 5340

[0534]	ggcggactta ttttgtttta tttcttagca tgcataagcaa tacatgattt gaggtttata	5400
[0535]	taataaaggg acaataaatc ttcaagaaac ttacccttac tgaattaaaa tattaagaa	5460
[0536]	ggtcacacat ttactcaaat atattagact actgggcaaa tagacatgaa aagtagagtt	5520
[0537]	aatattgagg taggccttct gtgaaatgtc taaggaaatt atgtttcata cagtgtgtaa	5580
[0538]	ccaagtggga atcatatcag aaagcagtca aaagcttata ttacaagtaa cagatgcttg	5640
[0539]	gttatatgac ctcccagagc ttgactgtct atacacaaaa agtgggtgta ataaaactgt	5700
[0540]	aatttgggct atgttttttt aaatggcttc accaacaatga aaggaaggga atgagcatgt	5760
[0541]	catggatgct tagagattat gcttccagca agaagaattg agctttggct cttattacag	5820
[0542]	aaacatgaca aggtgtgagt tttatttatt agaaattata taatatatta agctggggac	5880
[0543]	taaaaatfff attgaaacaa acaggcaagg gataggcatg tactagaagc aaaaatagga	5940
[0544]	tgtcaatgct gtaatgttat tttttggacc aaaatagtat ttcctataga aatgacaatg	6000
[0545]	atcttaggtt attattcttc ataaagatga caagttcaca agatatecta gttcattaaa	6060
[0546]	atcgttttag tcatttaata gagtgtctgt atagattaca caaaggaaag cacttacgat	6120
[0547]	gagaaataat gatatecaca attattttct taattcttag aaacattcta ttgttatatc	6180
[0548]	tcaatctcag aagccactta ttgctttatt attgaaacat atgaaattgt aagttatata	6240
[0549]	ttgtctatgg tgacatttca aagaacatgt gacgtacagt gtagcacaga taaagaacat	6300
[0550]	aactgcagct gaatcagtaa ctaaacttac atacattaaa tctgccatgt tggcaacagt	6360
[0551]	gtgtgcacta ccaaaggatg tactaatgct cagcacactc ccctatgtca cctttgttc	6420
[0552]	atcattacat cataggctca ttttgtttgc ttttgaaatc tagaccaagt cttttgtgtc	6480
[0553]	tttccaagca cagagctcat taatttacct catagacttg ttaaacttct tctggttcat	6540
[0554]	caattgaata gaaatactca ctactaatta tgtgagaccc tgccagtacc atagcacatg	6600
[0555]	gataatffff acataaaaca tgcatacaag taagattatt cagactgaac atgaatttta	6660
[0556]	gagaaatcag gaaggagtat atgggagtgg ttggagttag actagagaaa tgtaattaaa	6720
[0557]	ctataatctc aatacaaaga tctactaagc aaaaaacatg aaacattgtc attcaagtga	6780
[0558]	aacatcagtc ttcaaattgg aaagatattt ttactaggaa aatgtctggt agatggttat	6840
[0559]	tatctagaaa acacaaaaat tagaaaacgg taaactttaa taaaaagaat aatacaatga	6900
[0560]	gactacatga aaagtcttta actaatgaaa caaataatctt gaaacttttt tcttaaaagt	6960
[0561]	ttaatatcaa taaccatcat ggaaattcaa attaaaacta tttacatatt acccctgaaa	7020
[0562]	taataactaa tacccaataa aaataatata acaaaaaaat ggcaatgcat gccatcatgg	7080
[0563]	atttgggaga gagaatgttc attgcagttc tgaatggata ctggtgccac cacggtgaaa	7140
[0564]	atctctgtat aggtccttcc aaaagctgaa aatagacata tcacaagacc tgccacacat	7200
[0565]	ttttcaagca aatacceaaa ggactctacc tgactgcaga gacactttct cataaaatat	7260
[0566]	tattgttgat ctattcataa tatctggaaa atagaaacag ccaagatgcc catcaactga	7320
[0567]	ttaatagatg ataaaattat tgtacatttc agtgaatat tattcagttt ttaagaaaaa	7380
[0568]	tgaaattatg taataagcat gtaaatggat atatcttgaa acaaccattc cccattatat	7440
[0569]	tacctaaca ttgaaagtcc aaaatcatat gatcttttta gtggatctac taatcttttg	7500

[0570]	ctatatgtat tttattgaac tacccatgga tgtgagataa ttggttaacaa cagcacatgg	7560
[0571]	gagagcatgg gatcattcaa ggaagattag agagaatgca ttttttagga gataatggag	7620
[0572]	gagcaataga aaggattaaa tgaggttact gatgaaagtg atggttagag aaggcaatat	7680
[0573]	gaggagggat aactagcact tagggccttt tgaaaaagac atagagaaaa tactattgta	7740
[0574]	gaaacttcct ataattgggtg tatagttata tacaccaaag agctcagatg gagttaccct	7800
[0575]	ataatggaaa tattaactac tttttatcac tgtgataaaa catcctgaac agagcaacat	7860
[0576]	agattgggaa gcatttactt tggcttacag ttctaacggg ataaaaattc atgatgaaag	7920
[0577]	aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga agcaggtgag agctcacatc	7980
[0578]	ttgaagtgtg agaatgtagc agagagaaca aactgcaaat gaccagaaaa tgcttttgga	8040
[0579]	tcagagccca taccctctg actgacttct ccagaaattc tgaacaaata aaactcccca	8100
[0580]	aacagagcca taactgaagg tccagtgtct gagactacta ggggtatttc ttattcaaac	8160
[0581]	cactacaatg ggggtggggg agcaatctc caagtaggca ctacacacag acaataaaaa	8220
[0582]	actctagtaa ctggaatgga ttgacttatt tgaattactt gccagtggag ctacatagag	8280
[0583]	cacaattatt gtatttaaat taccctttat gatcttaca aacttgacag taagatcata	8340
[0584]	ttgctaaaga aaccacatat ttgaatcagg gaacatggtg atatctagtt gtcttcaac	8400
[0585]	tggaaacttc atgctttctg cccagcattc atgttgctgg aaagagcaat gtacactacc	8460
[0586]	agtgtagaaa ttaaatcatc aatcttatca agatgtggat cctataagtt acaataaaaa	8520
[0587]	ttagcctgat aagatatccc caccagaaga atattcacat aaatgctatg ggagcaacaa	8580
[0588]	gctattttct aaattagctt taatcctatt ctacaagaga gaatccatat ctagaatagt	8640
[0589]	tatagggatc aagaacccat ggcttgattg gtcataggcc caatgggaga tcctaataatt	8700
[0590]	attgttctac aaaatgaaaa taactcctaa tgacttggtg ctgcagtaat aagttagtat	8760
[0591]	gttgctcaac tctcacaaga gaagttttgt cttacaataa atggcaatta aagcagcccc	8820
[0592]	acaagattta tatcataccg atctctcat ggcctatgca tctagaagct aggaaacaaa	8880
[0593]	gaggacccta agagagacat acatggtccc cctggagaag gggaaggggg caagacctcc	8940
[0594]	aaagctaatt gggagcatgg gggaggggag agggagttag aagaaagaga aggggataaa	9000
[0595]	aggagggaga ggaggacaag agagagaagg aagatctagt caagagaaga tagaggagag	9060
[0596]	caagaaaaga gataccatag tagagggagc cttgtatgtt taaatagaaa actggcacta	9120
[0597]	gggaattgtc caaagatcca caaggtccaa ctaataatct aagcaatagt cgagaggcta	9180
[0598]	ccttaaaagc ctttctctga taatgagatt gatgactacc ttatatacca tcctagagcc	9240
[0599]	ttcatccagt agctgatgga agcagaagca gacatctaca gctaaacact gagctagtgtg	9300
[0600]	cagacaggga ggagtgatga gcaaagtcaa gaccaggctg gagaacaca cagaaacagc	9360
[0601]	agacctgaaa aaaaattgtc acatggaccc cagactgata gctgggagtc cagcatagga	9420
[0602]	cttttctaga aaccctgaat gaggatatca gtttgagggt ctggttaatc tatggggaca	9480
[0603]	ctggtagtgg atcaatatat atccctagtt catgactgga atttgggtac ccattccaca	9540
[0604]	tggaggaatt ctctgtcagc ctagacacat gggggagggt ctaggtcctg ctccaaataa	9600
[0605]	tgtgttagac ttgaagaac tcccttgaga agactcacc tccctgggga gcagaaaggg	9660

[0606] gatgggatga ggggttggga gggacaggag aggaggggag ggtgaggga ctgggattga 9720

[0607] caagtaaatg atgcttgttt ctaattttaa tgaataaagg aaaagtaaaa gaagaaaaga 9780

[0608] aaacaggcca aaagattata aaagacagag gtgggtgggtg actataaaga aacactatta 9840

[0609] tctaaataaa aatatgtcag aagcacacat gaacttatag tgtttatgaa agtatgtata 9900

[0610] ataactacat aatctcaagc caagaaaaaa atatcatctt tcagtgatga aggtgatattt 9960

[0611] atttctccca gaattaaagc caaagacctt atgaaagtaa ttatcttcaa aaggttgaaa 10020

[0612] atacatactt tgcaatacac agatctgcct agaaatctca tgttcacaat acacatgatg 10080

[0613] ctcaattgaa ttccattcaa tgttacagtt tagataaaca gttttagat aaactcacia 10140

[0614] tgtatcattt ctttttattt ttgacccaaa cagcttctca tctgttattc agaataattc 10200

[0615] ctcatggca ggatatccat cccaattggg ggaaggggag aatttgaaga aaacctagac 10260

[0616] cacatacata ttgccattg ggaacaaaag tctaaaatga tgtgttcac atcttctcta 10320

[0617] ctagtccctt ccccgctcca aagaaccttg gtatatgtgc ctcatcttct agagagagga 10380

[0618] aagcaggaac tgagcatccc ttacttgcca tctcaaccc aaaatttgca tcattgctca 10440

[0619] gctctgccct tctcatatga cagttacaag tcaaggcttc caaagtcct ctgtcatgtt 10500

[0620] tgggtgcaat agtttataca gatgacttca tgtcttcata tctaattgtt tatatagatt 10560

[0621] aatattaaac aatgttattt ctctaaccac attttaaatt aatttaaaaa tccattaatt 10620

[0622] gtgtctataa aatgcagaca gagtgtgag acacaatata agcctgatga tctgaatttg 10680

[0623] aaactcacac ccaccacatg gagaatcaac ttccaaaaat tttctatta ctccacact 10740

[0624] tacaccattg tacaacaca ataataatga acaaaatgaa atgaaataaa aaattaagtc 10800

[0625] tctgtaggta atgctactgt gcagcaaaag taaaaatggc agcttaagct tgctttatgg 10860

[0626] ttacacttta ccatcttcca ttaattataa ggacttcaat catggcagaa ctatgctgtt 10920

[0627] attgtctcag tgtaacctaa ccagggtgtc cagatgttct taatgtggac acctaaacta 10980

[0628] tttgatattt ggggttaagat ctttccctct ttcagaagaa acctcaggac agagggaatc 11040

[0629] ttgtctttta attttgagtc tgtagacttt ttccatttca aatatacatg aaacaagtga 11100

[0630] tgaagaaaat taatcaaaag gtgggaattg caatgatatt aggttcaata ttaagcttca 11160

[0631] atattatcat ggaatgcctt gttatacact gagtgtttgg caataaggga tttttagaag 11220

[0632] aaggagtttt tattctcaac aggttcttca agtttagctc aaataaatct aagcaatcca 11280

[0633] ctctagaatt aaatagtttc ctaagggcac agctatgaat agagctcaat ttacatataa 11340

[0634] aattttgttc accatttatg tcattccagt ttccattagt acaaggaaaa tacaaaatat 11400

[0635] ttagatgtca atatcaagtg aatagttcat ctctttttt aatatatac acctaaatca 11460

[0636] ccattttctc agaaaaatct ggcctgaagt tctgtctgga acttcaacat gaaaaatatg 11520

[0637] cacagcttgc tattataaat cctagttgat ttttaagatt catgtctggt gtctgactca 11580

[0638] gaggggccag aggctagaca aatatTTTTT gaatcttcat tgtgaagatt tttaatgatt 11640

[0639] attttaatat aaataacaaa gatgatggat aatgtaactt tgtacagttc atagacgctg 11700

[0640] aactactttg tgcttaaaat gttagtcccc tatcataaat gataggatgat aagtgtatgt 11760

[0641] ttaatacttt ccctctgagc tatattcatg tactagagaa ttattttaaa catgaaaaga 11820

[0642] ctgtgtttat agtctcagct cctgagaact ggtccaacct taggcagggtg aatgccagga 11880

[0643] gcaacgtttt tcttctacag aggatgcttt gctgccaagc aacctgggtg tgtggaaatg 11940

[0644] ttcctttttt aatcaagttt aaagggtctt catcatgctg ttgctccaca tattttcagg 12000

[0645] ttagagcttg gtccctggag tattatcttt taccagaaaa ttcatagtat tctttcaata 12060

[0646] actaacaact aaacttttcg ataaaaaaga attggaattt caatttttaa gcctgagtaa 12120

[0647] aattcttgtg aatcaggata ttttatttta agtcttatct tttaaaaagt tattttattt 12180

[0648] tttaaaaat tataatatat tttcataatt tcctccttc acttttctt acaaactt 12240

[0649] ctatagatca ccatgtgtt ttttttttac atttatggcc tctttctgtt cattgttatt 12300

[0650] acatacaaat agtcttgcc atagaagaac accacaattt gttacctgat aacaaattat 12360

[0651] caacccttaa aacctacaaa ctattgatat tactgaaaag actatactta tagatgtaaa 12420

[0652] gatatatgtg tgtgcacata tatagataca catatatgta ggatttttaa ttttagattt 12480

[0653] tagacatcaa aattatttat atgactgaga aactagacac tataaatgag cattcagtat 12540

[0654] tcaacaccgt gattttagat attgicacaa tgacagaaaa tttcttata gaaaatttta 12600

[0655] agttttgtga ttgctctgtg cacttagtga agtctcacag aaaaagaatc atagtatttt 12660

[0656] tagtttataa taaaaagtac atataattaa aatggttggc acaaaacaac atttgagcat 12720

[0657] ttttctatt tactatcaag tagtatcatt ttgaaataat aatttgacta gtttcaaaaa 12780

[0658] tgaaaacaaa atttaacta aatgccta atagcctgat aacattttta tgaatgaaat 12840

[0659] tattcaatag tgttatcaat taggggcccc aaacttttcc taaaataaaa cttttaattt 12900

[0660] ttttccattt ttatttaa atagaacaaa attgttttac atgtaaatca gagtttctc 12960

[0661] accctccct tctcctgtc cctcactaac accctacttg tccatacca tttctgtcc 13020

[0662] ccaggaggagg tgaggccttc catggggaaa cttcagagtc tgtctatcct ttcggatagg 13080

[0663] gcctaggccc tcaccattt gtctaggcta aggctcacia agtttactcc tatgctagtg 13140

[0664] ataagtaact atctactaca agagacacca tagatttctt aggttcttc actgacaccc 13200

[0665] atgttcatgg ggcttggaac aatcatatgc tagtttctta ggtatcagtc tggggacccat 13260

[0666] gagtcccc ttgttcaggt caactgtttc tgtgggtttc accaccctgg tcttgactgc 13320

[0667] ttgtctatc actcctccct tctgttaact gggttccagt acaattccgt gtttagctgt 13380

[0668] ggggtgtctac ttctacttc atcagcttct gggatggagc ctctaggata gcatacaatt 13440

[0669] agtcatcctc tcattatcag ggaaggcat ttaaagtagc ctctccattg ttgcttggat 13500

[0670] tgtagttgg tgtcatcttt gtagatctct ggacatttcc ctagtccag atatctcttt 13560

[0671] aaacctacaa gactacctt attatggat ctcttttctt gctctcgtct attcttccag 13620

[0672] acaaatctt cctgtccct tataatttcc tctccctcc tcttctccc ttctcattct 13680

[0673] cctagatcca ttttccctc ccccatgtc ccaagagaga tgttgctcag gagatcttgt 13740

[0674] tccttaacc ttttcttggg gatctgtctc tcttagggtt gtccttgttt cctagcttct 13800

[0675] ctggaagtgt ggattgtaag ctggtaatca tttgtccat gtctaaaatc catatatgag 13860

[0676] tgatgtttgt ctttttgtga ctgggttacc tcttcaaaa tggtttcttc catatgtctg 13920

[0677] tggatttcaa tagcacaac aacatacag atcttggggc aactaacc aaacaagtga 13980

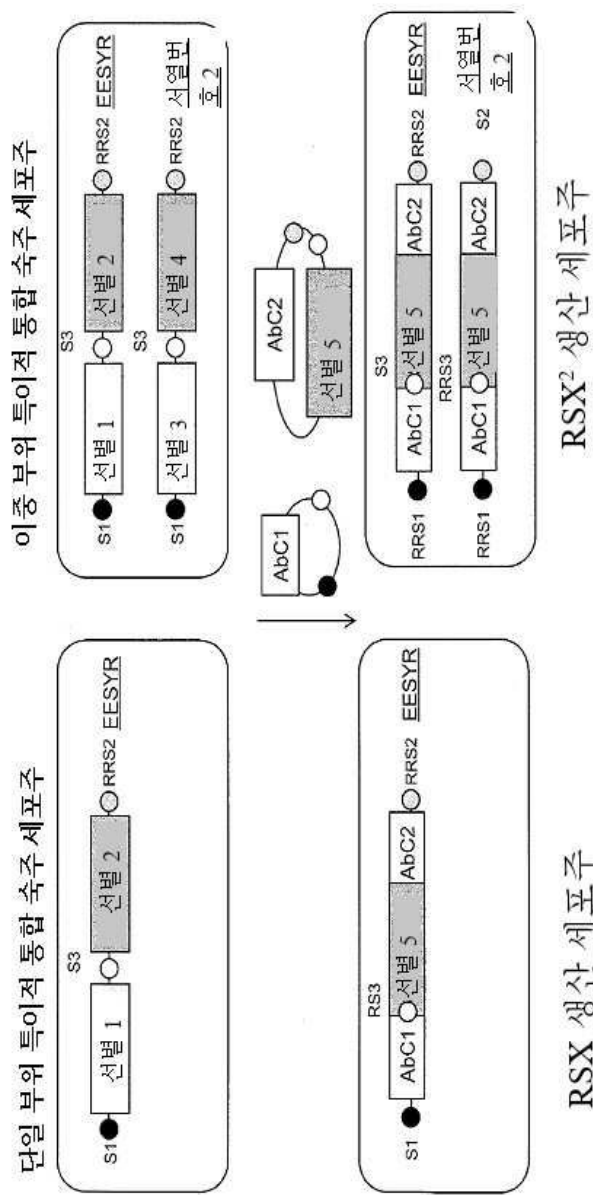
[0678]	aagaccagta tagcaagaac tttagagttta aagaaagaaa ttaaagaaga taccagaaaa	14040
[0679]	tggaaagatc tcccatgctc tttagataggc agaatacaaca tagtaaaaaat ggcaatcttg	14100
[0680]	ccaaaatcca tctacagact caatgcaatc cccattaaat accagcacac ttcttcacag	14160
[0681]	acctgaaaga ataatactta actttatattg gagaaacaaa agaccagga taggccaac	14220
[0682]	aacctgtac aatgaaggca cttccagagg catcccatc cctgacttca agctctatta	14280
[0683]	tagagtaata atcctgaaaa cagcttggta atggcacaaa aatagacagg tagaccaatg	14340
[0684]	gaattgagtt gaaaacctg atattaaccc acatatctat gaacacctga ctttgacaaa	14400
[0685]	gaagctaagg ttatacaatg taagaaagaa agcatcttca acaaatcgtg ctggcataac	14460
[0686]	tggatgctgg catgtagaag actgcagata gatccatgct taatgccatg cacaaaactt	14520
[0687]	aagtccaaat ggatcaaaaa cctcaacata aatccagcca cactgaacct catagaagag	14580
[0688]	aaagtgggaa gtatccttga ataaattggt acaggagacc acatcttgaa cttaacacca	14640
[0689]	gtagcacaga caatcagatc aataatcaat aaatgggacc tctgaaact gagaagcttc	14700
[0690]	tgtaggcaa tggataagtc aacaggacaa aatggcagcc cacggaatgg gaaaagatat	14760
[0691]	tcaccaatcc tatatctgac agagggctgc tctctatttg caaagaacac aataagctag	14820
[0692]	tttttaaac accaattaat ccgattataa agttgggtag agaactaaat aaagaattgt	14880
[0693]	taacagagca atctaactg gcagaaagac acataagaaa gtgctcacca t	14931
[0694]	서열번호 3	
[0695]	4001 염기	
[0696]	DNA	
[0697]	<i>Cricetulus griseus</i>	
[0698]	ccaagatgcc catcaactga ttaatagatg ataaaattat tgtacatttc agtgtaatat	60
[0699]	tattcagttt ttaagaaaa tgaaattatg taataagcat gtaaatggat atatcttgaa	120
[0700]	acaaccattc cccattatat tacctaaca ttgaaagtcc aaaatcatat gatcttttta	180
[0701]	gtggatctac taatcttttg ctatatgtat ttatttgaac taccatgga tgtgagataa	240
[0702]	ttggtaacaa cagcacatgg gagagcatgg gatcattcaa ggaagattag agagaatgca	300
[0703]	ttttttagga gataatggag gagcaataga aaggattaaa tgaggttact gatgaaagtg	360
[0704]	atggtttagag aaggcaatat gagggaggat aactagcact tagggccttt tgaaaaagac	420
[0705]	atagagaaaa tactatttga gaaacttcct ataattgggtg tatagttata tacaccaaag	480
[0706]	agctcagatg gagttaccct ataattgaaa tattaactac tttttatcac tgtgataaaa	540
[0707]	catcctgaac agagcaacat agattgggaa gcatttactt tggcttacag ttctaacggg	600
[0708]	ataaaaattc atgatgaaag aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga	660
[0709]	agcaggtgag agctcacatc ttgaagtgtg agaattgtgc agagagaaca aactgcaaat	720
[0710]	gaccagaaaa tgcttttggg tcagagccca taccctctg actgacttct ccagaaattc	780
[0711]	tgaacaaata aaactcccca aacagagcca taactgaagg tccagtgtct gagactacta	840
[0712]	ggggtatttc ttattcaaac cactacaatg ggggtggggg agcaatcctc caagtaggca	900
[0713]	ctacacacag acaataaaaa actctagtaa ctggaatgga ttgacttatt tgaattactt	960

[0714]	gccagtggag ctacatagag cacaattatt gtattttaa	1020
[0715]	aacttgacag taagatcata ttgctaaga aaccacatat	1080
[0716]	atatctagtt gttcttcaac tggaaacttc atgctttctg	1140
[0717]	aaagagcaat gtacactacc agtgtagaaa ttaaatcatc	1200
[0718]	cctataagtt acaataaaaa ttagcctgat aagatatccc	1260
[0719]	aaatgctatg ggagcaacaa gctattttct aaattagctt	1320
[0720]	gaatccatat ctagaatagt tatagggatc aagaacccat	1380
[0721]	caatgggaga tcctaattatt attgttctac aaaatgaaaa	1440
[0722]	ctgcagtaat aagttagtat gttgctcaac tctcacaaga	1500
[0723]	atggcaatta aagcagcccc acaagattta tatcataccg	1560
[0724]	tctagaagct aggaaacaaa gaggacccta agagagacat	1620
[0725]	gggaaggggg caagacctcc aaagctaatt gggagcatgg	1680
[0726]	aagaaagaga aggggataaa aggagggaga ggaggacaag	1740
[0727]	caagagaaga tagaggagag caagaaaaga gataccatag	1800
[0728]	taaatagaaa actggcacta gggaattgtc caaagatcca	1860
[0729]	aagcaatagt cgagaggcta ccttaaaagc ctttctctga	1920
[0730]	ttatatacca tcctagagcc ttcatccagt agctgatgga	1980
[0731]	gctaaacact gagctagttg cagacaggga ggagtgatga	2040
[0732]	gagaaacaca cagaaacagc agacctgaaa aaaatgttgc	2100
[0733]	gctgggagtc cagcatagga cttttctaga aacctgaat	2160
[0734]	ctggttaatc tatggggaca ctggtagtgg atcaatat	2220
[0735]	atttgggtac ccattccaca tggaggaatt ctcgtgcagc	2280
[0736]	ctaggtcctg ctccaaataa tgtgttagac ttgaagaac	2340
[0737]	tccttgggga gcagaaaggg gatgggatga gggttggga	2400
[0738]	ggtgagggaa ctgggattga caagtaaatg atgcttgttt	2460
[0739]	aaaagtaaaa gaagaaaaga aaacaggcca aaagattata	2520
[0740]	actataaaga aacactatta tctaataaa aacatgtcag	2580
[0741]	tgtttatgaa agtatgtata ataactacat aatctcaagc	2640
[0742]	tcagtgatga aggtgatttt atttctccca gaattaaagc	2700
[0743]	ttatcttcaa aaggttgaaa atacatactt tgcaatacac	2760
[0744]	tgttcacaat acacatgatg ctcaattgaa ttccattcaa	2820
[0745]	gtttgtagat aaactcaca tgtatcattt ctttttattt	2880
[0746]	tctgttattc agaataattc ctcgatggca ggatatccat	2940
[0747]	aatttgaaga aaacctagac cacatacata ttgccattg	3000
[0748]	tgttgttcac atcttctcta ctagtctct ccccgtecca	3060
[0749]	ctcattttac agagagagga aagcaggaac tgagcatccc	3120

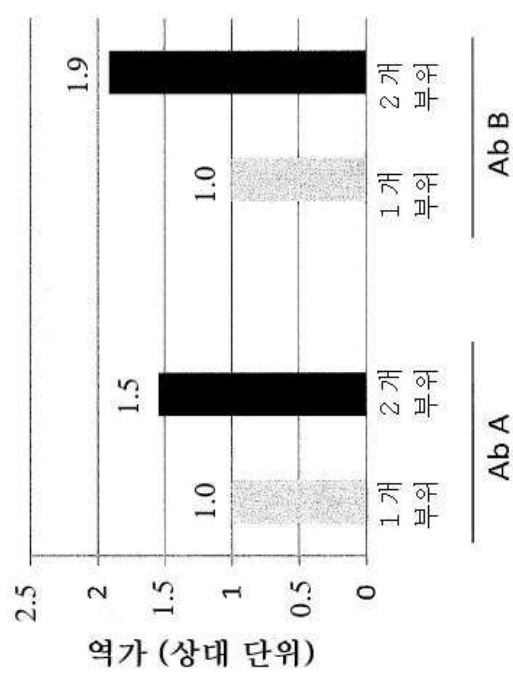
[0750]	aaaatttgca tcattgctca gctctgccct tctcatatga cagttacaag tcaaggcttc	3180
[0751]	caaagtcctt ctgtcatgtt tgggtgtcaat agtttataca gatgacttca tgtcttcata	3240
[0752]	tctaattgtct tatatagatt aatattaaac aatgttattt ctctaaccac attttaaatt	3300
[0753]	aatttaaaaa tccattaatt gtgtctataa aatgcagaca gagtgctgag acacaatata	3360
[0754]	agcctgatga tctgaatttg aaactcacac ccaccacatg gagaatcaac ttccaaaaat	3420
[0755]	tttctatta cttccacact tacaccattg tacaacaca ataataatga acaaaatgaa	3480
[0756]	atgaaataaa aaattaagtc tctgtaggta atgctactgt gcagcaaaag taaaaatggc	3540
[0757]	agcttaagct tgctttatgg ttacacttta ccatcttcca ttaattataa ggacttcaat	3600
[0758]	catggcagaa ctatgctgtt attgtctcag tgaacctaa ccagggtgtc cagatgttct	3660
[0759]	taatgtggac acctaaacta ttgatattt gggttaagat ctttcctctt ttcagaagaa	3720
[0760]	acctcaggac agaggggaatc ttgtctttta attttgagtc tgtagacttt ttccatttca	3780
[0761]	aatatacatg aaacaagtga tgaagaaaat taatcaaaag gtgggaattg caatgatatt	3840
[0762]	aggttcaata ttaagcttca atattatcat ggaatcgctt gttatacact gagtgtttgg	3900
[0763]	caataaggga tttttagaag aaggagtttt tattctcaac aggttcctta agtttagctc	3960
[0764]	aaataaatct aagcaatcca ctctagaatt aaatagtctc c	4001

도면

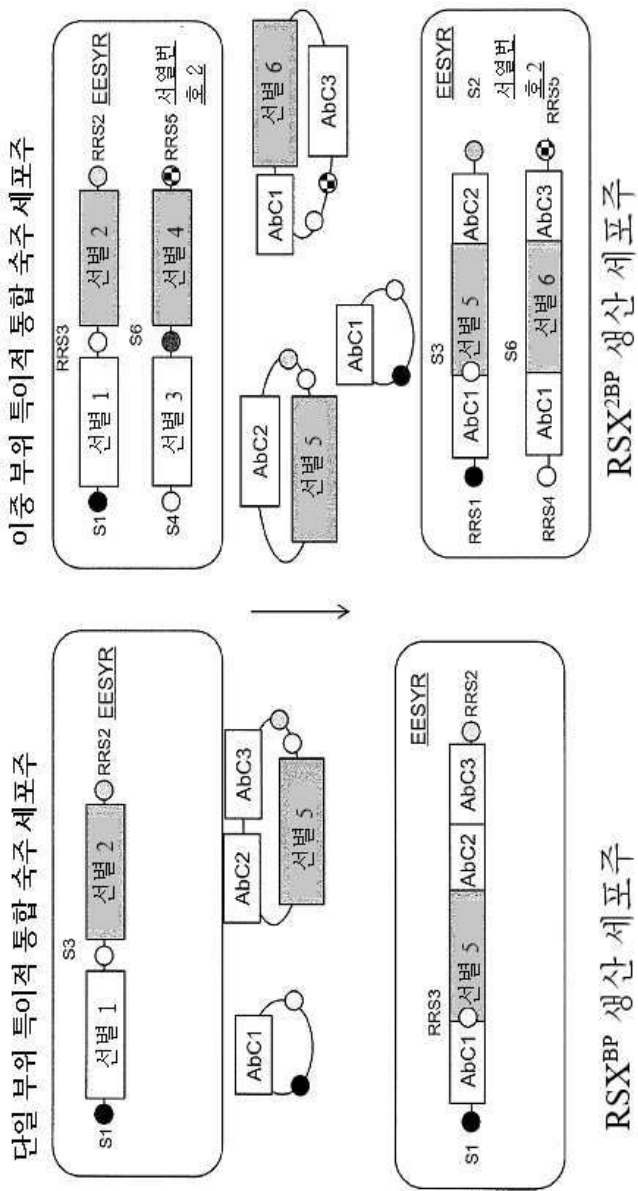
도면1



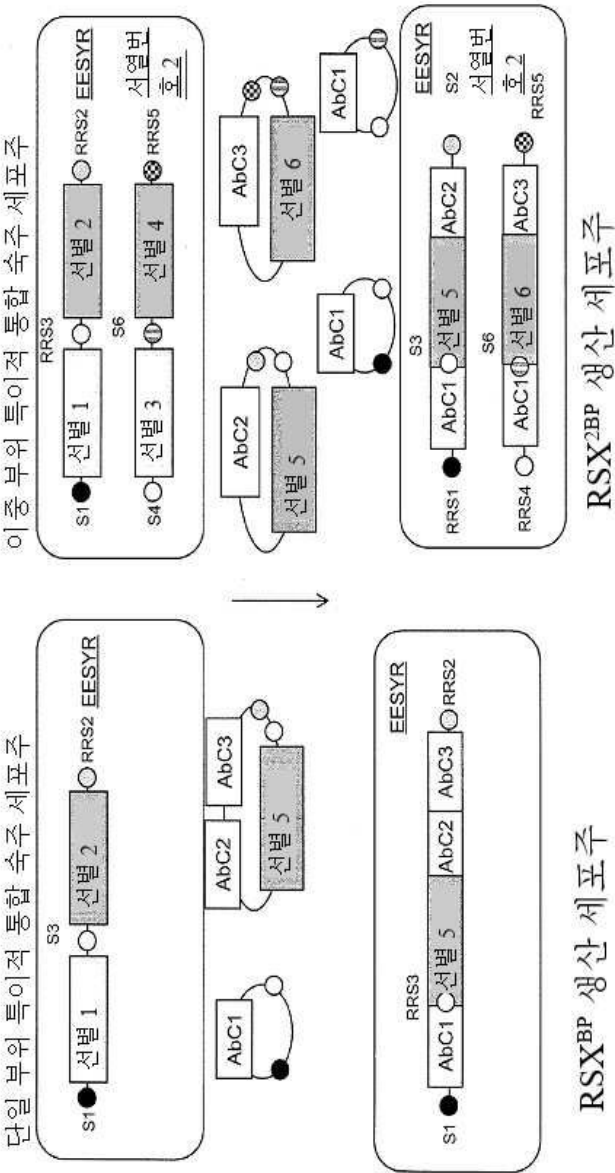
도면2



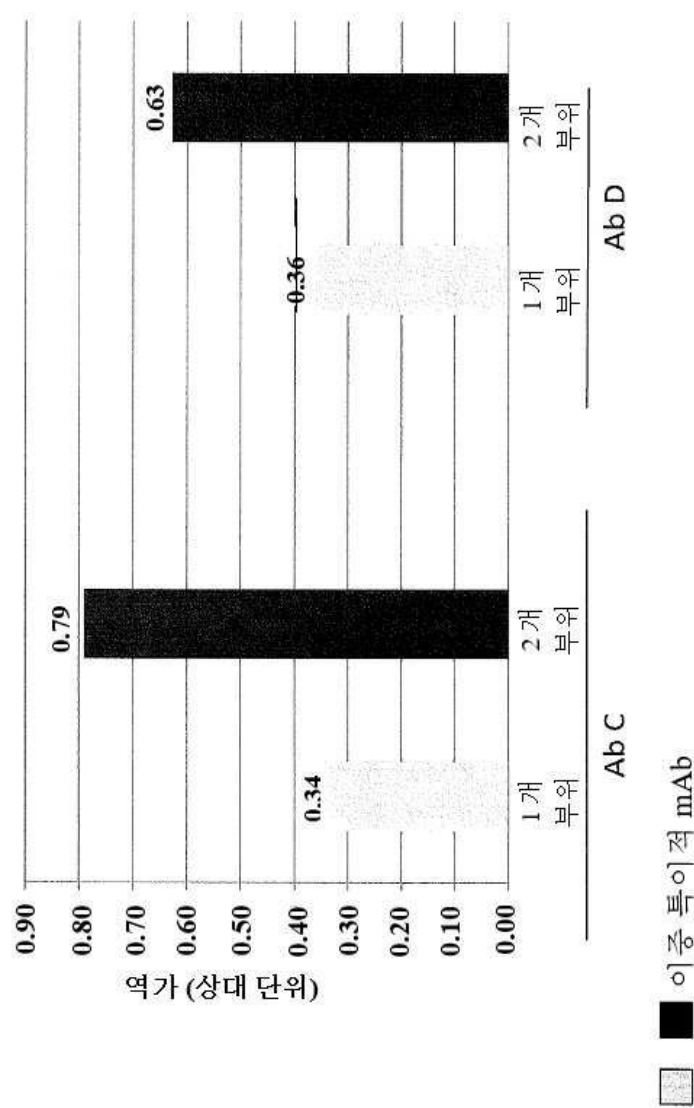
도면3



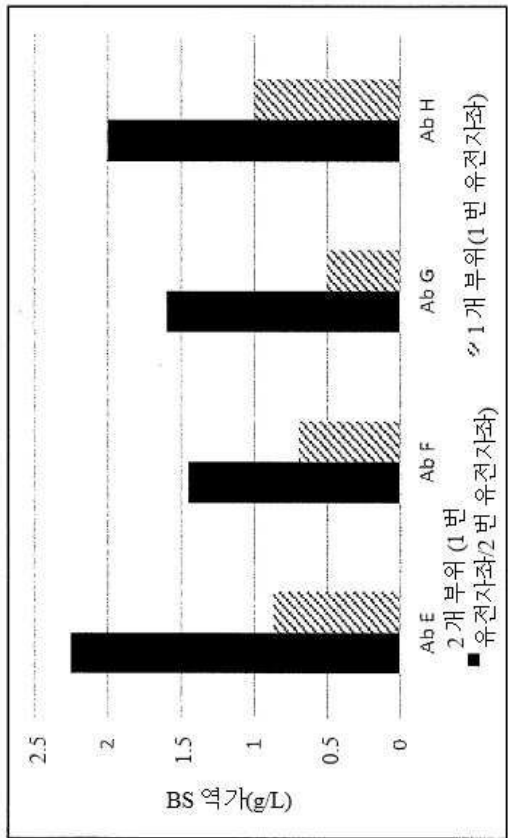
도면4



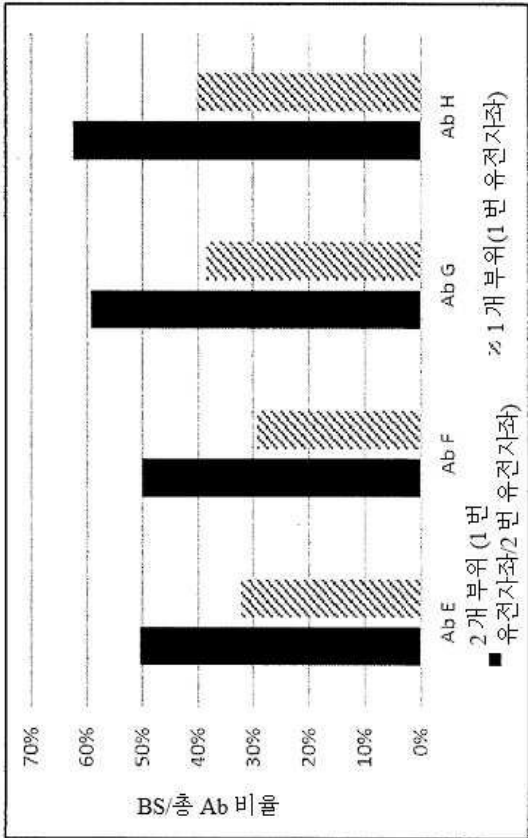
도면5



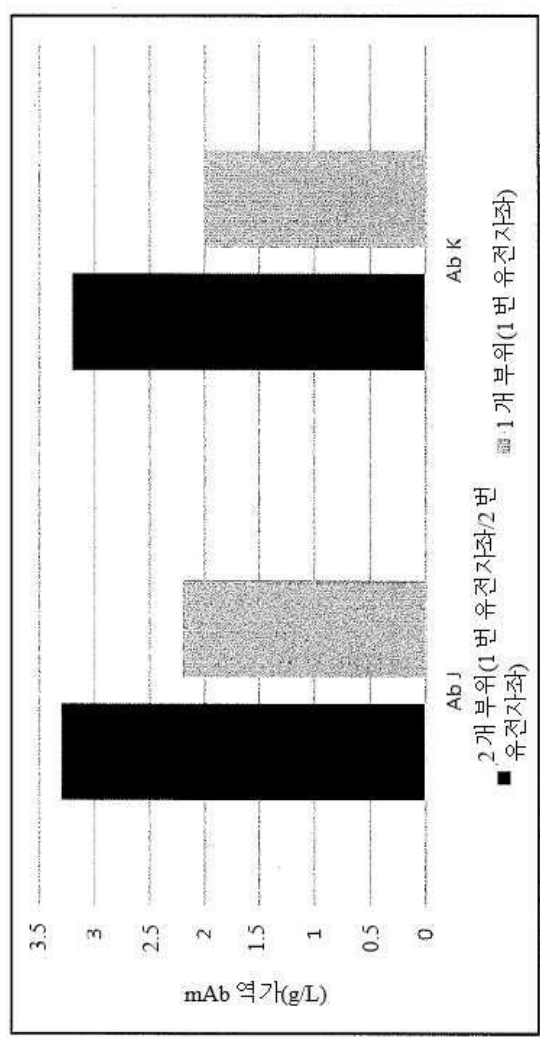
도면6a



도면6b



도면7



서열 목록

SEQUENCE LISTING

- <110> Regeneron Pharmaceuticals, Inc.
- <120> COMPOSITIONS AND METHODS FOR MAKING ANTIBODIES BASED ON USE OF
EXPRESSION-ENHANCING LOCI
- <130> 33444PCT (10341W001)
- <150> 62/325,400
- <151> 2016-04-20
- <160> 3
- <170> PatentIn version 3.5
- <210> 1
- <211> 13515
- <212> DNA

<213> Cricetulus griseus

<400> 1

tctagaaaca aaaccaaaaa tattaagtca ggcttggtt caggtgctgg ggtggagtgc	60
tgacaaaaat acacaaattc ctggctttct aaggcttttt cggggattca ggtattgggt	120
gatggtagaa taaaaatctg aaacataggt gatgtatctg ccatactgca tgggtgtgta	180
tgtgtgtgta tgttgtctg tgttgtgcc cagacagaaa taccatgaag gaaaaaaca	240
cttcaagac aggagagaag agtgacctgg gaaggactcc ccaatgagat gagaactgag	300
cacatgccag aggaggtgag gactgaacca ttcaacacaa gtggtgaata gtcctgcaga	360
cacagagagg gccagaagca ctcagaactc caggggttca ggagtgggtc tctggaggct	420
tctgcccttg gaggttctg aggaggaggc ttccatattg aaaatgtagt tagtgccgt	480
ttccattagt acagtgacta gagagagctg agggaccact ggactgaggc ctagatgctc	540
agtcagatgg ccatgaaagc ctagacaagc acttccgggt ggaaaggaaa cagcaggtgt	600
gaggggtcag gggcaagtta gtgggagagg tcttcagat gaagtagcag gaacggagac	660
gcactggatg gccccacttg tcaaccagca aaagcttga tctgttcta agaggccagg	720
gacatgacaa gggtagctc ggtttttaa aggctttgtg ttacctaact acttctatta	780
gtcagatact ttgtaacaca aatgagtact tggcctgtat tttagaaact tctgggatcc	840
tgaaaaaaca caatgacatt ctggctgcaa cacctggaga ctcccagcca ggccctggac	900
ccgggtccat tcatgaaat actcaggac agattcttca ctaggtactg atgagctgctc	960
ttggatgcaa atgtggctc ttcattttac tacaagtcac catgagtcag gagtgctgt	1020
ttgcacagt tgactaagtg atggagtgtt gactgcagcc attcccgcc ccagcttgtg	1080
agagagatcc ttttaaatg aaagtaagct caaagttacc acgaagccac acatgtataa	1140
actgtgtgaa taatctgtgc acatacacia accatgtgaa taatctgtgt acatgtataa	1200
actgtgtgaa taatctgtgt gcagcctttc cttacctact accttccagt gatcaggttt	1260
ggactgcctg tgtgctactg gacctgaat gtccccaccg ctgtccctg tcttttacga	1320
ttctgacatt ttttaataat tcagcggtt ccctctgct ctgtgcctag ctataccttg	1380
gtactctgca ttttggttc tgtgacattt ctctgtgact ctgctacatt ctcagatgac	1440
atgtgacaca gaagggttc cctctggaga catgtgatgt cctgtcatt agtggaatca	1500
gatgccccca aactgtgtc cagtgtttgg gaaagtgaca cgtgaaggag gatcaggaaa	1560
agaggggtgg aaatcaagat gtgtctgagt atctcatgct cctgagtggt ccaggctgct	1620
gacttcactc cccaagtga gggaggccat ggtgagtaca cacacctac acatactata	1680
tccaacacac acacacacac acacacacac acgcacgcac gcacgcacgc acgcacacat	1740

gcacacacac gaactacatt tcacaaacca catacgcata ttacacccca aacgtatcac 1800

ctatacatat cacacatata caccctcca cacatcacac acataccaca cccacacaca 1860

gcacacacat acataggcac acattcacac accacacata tacatttgtg tatgcataca 1920

tgcatacaca cacaggcaca cagacaccac acacatgcat tgtgtacgca cacatgcata 1980

cacacacata ggcacacatt gagcacacac atacatttgt gtacgcacac tacatagaca 2040

tatatgcatt tgtatatgca cacatgcatg cacacatata taggcacaca tagagcacac 2100

acatacatatt gtgtatgcac acatgcacac accaatcaca tgggaagact caggttcttc 2160

actaaggttc acatgaactt agcagttcct ggttatctcg tgaaacttgg aagattgctg 2220

tggagaagag gaagcgttgg cttgagccct ggcagcaatt aaccccgccc agaagaagta 2280

ggtttaaaaa tgagagggtc tcaatgtgga acccgaggg cgccagttca gagaagagac 2340

ctaccaagc caactgagag caaaggcaga gggatgaacc tgggatgtag ttgaacctc 2400

tgtaccagct gggcttcatg ctattttgtt atatctttat taaatattct tttagtttta 2460

tgtgcgtgaa taccttgctt gcataaatgt atgggcactg tatgtgttct tggtgccggt 2520

ggaggccagg agagggcagtg gatcctccgg agctggcgtt tgagacagtt gtgaccaca 2580

gtgtggggtc tgggaactgg gtcttagtgi tccgcaagtg cagctggggc tcttaacctc 2640

tgagccatcc ctccagcttc aagaaactta tttcttagg acatggggga agggatccag 2700

ggcttttagc ttgtttgttc agcaaatact cttttcgtgt attttgaatt ttattttatt 2760

ttactttttt gggatagaat cacattctgc agctcaggct gggcctgaac tcatcaaat 2820

cctcctgtct cagtctacca ggtgataaga ttactgatgt gagcctggct ttgacaagca 2880

ctttagagtc cccagccctt ctggacacit gtccaagta taatatatat atatatatat 2940

atatatatat atatatatat atatatgtg tgtgtgtgtt tgtgtgtgta tgagacactt 3000

gctctaagggt tatcatatat atccttgatt tgcctttaat ttatttttta attaaaaatg 3060

attagctaca tgcacctgt atgcgtctgt atcatctata tatccttctc tccttctctc 3120

tccttctctc ttcttctctc cacccecaag catctatctt caaatccttg tgccgaggag 3180

atgccaagag tctcgttggg ggagatgggt agggggcgat acaggggaag agcaggagga 3240

aagggggaca gactgggtgt ggtctttgga gagctcagga gaatagcagc gatcttcct 3300

gtccctgggt tcacctctta cagccaacac cattttgtgg cctggcagaa gattgtcaa 3360

gctgggtcga ggctgcccac acaaccccaa tctggcccca agaaaaggca cctgtgtgtg 3420

actctggggt taaaggcgt gcctgggtcgt ctccagctgg acttgaaact cccgtttaat 3480

aaagagtctt gcaaaataat acccgagag tcacagtgcc aggttcccgt gctttcctga 3540
 agcgccaggc acgggttccc taggaaatgg ggccttgctt gccaaagtcc caccgcttgc 3600
 cctgcaaacg gcctgaatga tctggcactc tgcgttgcca ctgggatgaa atggaaaaaa 3660
 gaaaaagaag aagtgtctct ggaagcgggc gcgctcacac aaacccgcaa cgatttgtga 3720
 aacactctcc attgagaatc tggagtgcgg ttgccctcta ctggggagct gaagacagct 3780
 agtggggcg gggggaggac cgtgctagca tcctccacg gtgctcgtg gctgtggtgc 3840
 atgccgggaa ccgaaacgcg gaactaaagt caagtcttgc tttggtggaa ctgacaatca 3900

 acgaaatcac ttcgattgtt ttctctttt tactggaatt cttggatttg atagatgggg 3960
 gaggatcaga gggggagggg agggcgggg agacggaggg aggaggggag gaggggagga 4020
 ggggaggagg ggaggagggg aagggatgga ggaaaatact aacttttcta attcaacatg 4080
 acaaagattc ggagaaagtg caccgctagt gaccggagg aggaatgccc tattgggcat 4140
 tatattccct gtcgtctaata ggaatcaaac tcttggttcc agcaccaagg attctgagcc 4200
 tatectattc aagacagtaa ctacagccca cacggaagag gctatacaac tgaagaaata 4260
 aaattttcac tttatttcat ttctgtgact gcatgttcac atgtagagag ccacctgtgt 4320

 ctaggggctg atgtgctggg cagtagagtt ctgagcccgt taactggaac aaccagaac 4380
 tcccaccaca gttagagctt gctgagagag ggaggccctt ggtgagattt ctttgtgtat 4440
 ttatttagag acagggtctc atactgtagt ccaagctagc ctccagctca cagaaattct 4500
 cctgttccgg ttccaaagt actggagtta tgagtgtgtg ttaattgaac gctaagaatt 4560
 tgctgattga agaaaacctc aagtgggttt ggctaatacc cagacccca gaggctgagg 4620
 caggaggaat gagagaattc aaggtttgcc agagccacag ggtgagctca atgtggagac 4680
 tgtgagggtg agctcaatgt ggagactgtg agggtgagct caatgtggag actgtgaggg 4740

 tgagctcaat gtggagactg tgagggtgag ctcaatgtgg agactgtgag ggtgagctca 4800
 atgtggagac ctgtatcaag ataataatag tagtagtaac aatgcaggcg aggggtgtgt 4860
 tgagtggtag agcagttagt tgatttgaca tgcttgaggt ctcccgttcc atctgtggcc 4920
 ctgcaacagg aaggaggagg ggaagggggg gaacgagaga gaggaagag agacagaagc 4980
 taagataggg aatgagagag gaaggaagaa acgggaagaa attcagactc cttcctgagt 5040
 tccccaacg ctagtgaca tctgtgcac accctaaggt ggcctttgtg tggcactggc 5100
 ttgggtggtc gggaaaggca ttttcagctt gttgcagaac tgccacagta gcatgctggg 5160

 tccgtgaaag tttctgcccg ttaacaagaa gtcttacta cttgtgacct caccagtga 5220
 aatttcttta attgtctcct ggtgttcttg gttttgcatt ttgttttcta aggatacatt 5280
 cctgggtgat gtcatgaagt ccccaaagac acagtggggc tgtgttgat tgggaagat 5340

gatttatctg ggggtgtcaaa aggaaaagaa gggaaacagg cacttgggaa aatgtcctcc 5400

cgccccccg aatttttggt tggcaaccgt ggtggaggag caagaaacac gtggacgttt 5460

gaggaggcat ggggtcctag gaggacagga agcagaagga gagagctggg ctgacagcct 5520

gcaggcattg cacagtttca gaaggagatt acagcatgac tgagttttta gggatccaac 5580

agggacctgg gtagagattc tgtgggctct gaggcaactt gacctagcc agatgttatt 5640

tgaataacct gctcttagag ggaaaacaga catagcaaac agagccacgt ttagtgatga 5700

aactctcact ttgcctgagt catgtgcggc catgcccagg ggtcaggctg aactcaact 5760

caaaaacaag tgagaaatg aagacaatcc gtggtggcag ctactggaag ggccaccaca 5820

tccccagaaa gagtggagct gctaaaaagc cattttgat aggcacagtt atcttgaatg 5880

catggagcag agattacgga aaaatcgaga atgttaatga ggcaacattc gatttgagtc 5940

attcagtgtg ggaaaccag acgcttccat ccctaaaag gaacatcttg ctctcagtca 6000

aaatggaaat aaaaattggg gcttgaattt ggcaaatgat tcagaactct gtgtaggtat 6060

tttcacacgc acagtggata atttcatgt tggagtttat ttgtgctaaa aggcagaaaa 6120

gggtaaaaag cacatcttaa gatttatgag gttctacgaa taaaaataat gttacttaca 6180

gctattcctt aattagtacc ccttccacc tgtggtaatt tctgagata gtcagtgggg 6240

aaaagatctc tcttctctt ctttctccc cteccctect cteccctect cctccctec 6300

ctccctecte tectecctc ccttctect tcttctttg ctecttctc tctgectect 6360

tctcccttc tcttcatctt attctaagta gcttttaaca gcacaccaat tacctgtgta 6420

taacgggaaa acacagctc aagcagctta gagaagattg atctgtgtc actagcgtgc 6480

aattcagagg tgggtgaaga taaaaggcaa acatttgagg ccatttctt atttggcacg 6540

gcacttagga agtggaaatc gcctaatact ctggtttgta ccactttcc ctataatgga 6600

ctgtttggga agctcctggg caaccgattc tggcatctca ttggtcagag gcctgttaaa 6660

tgggtactctt atttgcaaag aaggctgtaa ctgttagctt taaaagcctc tctcaagaa 6720

agaagggaga aaggatatgg ctagacatat ctaatagact taaccactgt gaaaagcctt 6780

agtatgaatc agatagaacc tatttttaac tcagttttga aaaaaataat ctttatattt 6840

atgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gaaccacatg 6900

tagcaggtgc tggaggaggc cagaagaggg caccagatct cctggaactg acaccacaca 6960

tggttatgag ctgcctgatg tgggtgctgg gaactgaact ctcgtgttct gcaagagcag 7020

caactgttct cttaactgat gagccatctc tccagcccc ccataattt taattgttca 7080

ttttagtaaa ttttattcat aatcaattat cacagtataa aacaatgatt ttatatatat 7140

catatacata tcaaggatga cagtgagggg gatatgtgtg tgtgtgtgtg tgtgtgtgtg 7200

tgtgtgtgtg tgtgttatit gtgtgtgtgc tttttaagaa ggtgccatag tcaactgcatt	7260
tctctgaagg atttcaaagg aatgagacat gtctgtctgc caggaaccct atcttcctct	7320
ttgggaatct gacccaaatg aggtattctg aggaactgaa tgaagagctc aagtagcagt	7380
gtcttaaac caaatgtgct gtctagagaa agtcaacgtc atcagtgagc tgaggagaga	7440
tttactgagc ggaagacaag cgctctttga ttttaagtggc tcgaacagtc acggctgtgg	7500
agtggagcct gtgctcaggt ctgaggcagt ctttgctagc cagctgtgat gagcagtga	7560
gaaaggttg agatggaggc aggggtgggag cagggtatg gttcagacta ggtatcgtga	7620
gcacaccagc tggttgactt gtggtctgtg ggtcaggcgt tgtaaacgcc ctcagggtca	7680
ggcagtcaca ttgcttgaag ctgaatgggt gaggcaacac agagagtga aagaaggcaa	7740
agtaccacct ctccccgac ccaggtcact tctgggttat agctgagact ccggacagca	7800
tgcaaccagc tggttagagc ttcaggga aa acttgatgtc tgcatgttgc tatgaaatgt	7860
gattcggtag atctggagaa aatttataat gctggctcag tcaagcactg aacaaaggta	7920
ccttggcttt gggagctaca tgacattgac ttgtaggcag actttttttt ttctgcccgc	7980
caattccag ataaccaata tggaggctca atattaatta taaatgctcg gctgatagct	8040
caggcttgtt actagctaac tctccaact taaatgaacc catttctatt atctacattc	8100
tgccacgtga ctttacctg tacttctgt ttcctctct tgtctgactc tgcccttctg	8160
cttcccagag tcttagtct gggttctctg cctaacctta tctgcccag ctgctgacca	8220
agcatttata attaatatta agtctcccag tgagactctc atccaggag gacttgggtg	8280
ctccccctc ctcatgcca tccgtgtctt cctcttccct cgttcccc tctcttctt	8340
gtcttctct ctcaccctt ctttcatag tattgatggc aagggtgttc tagaatggag	8400
gagtgtccat aggcagca agaaaccagt taggatgtc tgtgaggggt tgtaatcata	8460
agcgatggac acaattcaag ccacagagt aagacggaag gatgcactgt gctctagagc	8520
aacttctggg gcagaatcac aggggtgagt tctgacttga gggcgaagag gccacgagga	8580
agggagttag tttgtctgag ctagaagcta cgccccact cttggtagca gacctgcca	8640
caagcatgct ttgttaatca tgtgggatct gattttctc taaatctatg ttcaactctt	8700
aagaaaatgt gaattctcac attaaaattt agatatacgt cttttggtgg ggggggtgta	8760
aaaaatctc aagaatatgg atttctgggg gccggagaga tggctcagag gtttaagagaa	8820
ctggttgctc ttctagacat tctgagttca attcccagca accacatggt ggctcacaac	8880
catctgtaat gcgacctggt gccatcttct gacatgcatg gatacatgca ggcagaaagc	8940

tgtatacata gtaaatlgat aaatcttttt ttaaaaagag tatggattct gccgggtgtt	9000
ggtggcgcac gcctttaatc ccagcactct ggaggcagag gcaggtggat ctctgtgagt	9060
tcgagaccag cctggctctat aagagctagt tccaggacag cctccaaagc cacagagaaa	9120
ccctgtctcg aaaaaccaa aaaaaaaaaa aaaaaaaaaa aaaaaaaga gtatggattc	9180
taagaaagcc gtaacagctg gagctgtgta cggagttagc cgtgtgacta gaagaacaga	9240
cattcatgat gaaacacccc aggattttta cttagtatct agtttccatt gtgttttga	9300
gaccggctct taigtctcc aggttggcct caaactgctg atcttcccgc ctctacctct	9360
caagtcttg gactacttgg ctcataaaac agtttttgc gggctccctg aagttatggt	9420
tgtacaaacc gtgggggtca atatactcac ttgggcagag agagaaggtc tgaatcccag	9480
acaatgactg catctcagga cagttgggaa gaggacaatg gcagaaggac ttagaaaaga	9540
tagactggag ggtggaaaag cagcaggaac agagaaacaa aacaggaagc ttgctatcca	9600
gggccactct ggagtctgt ggcaaatgg aagcgggcta ggggaataca ttgtgtctac	9660
tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgat caatgcctat caatgttgaa	9720
ggggaaatat gtataccaca ttgattctgg gagcaattct cagtatctgg cctagagaaa	9780
ggaatggccc ctgcagaata gacagagtga atggtgcctt ttatcatttg cttaaagtga	9840
ggagaaataa acatccttcc atagagtttc aggtaaatga accccacagt tcatctgtgc	9900
cgtgggtggag gcctggccaa cagttaaaaa gattagacac ggacaaagtc tgaaggaaac	9960
acctgaata ggaagaggag agccacctca ttctgtaact ttctcaagg ggaagatgtt	10020
ccaagagtgg gaataaatgg tcaaaggggg gatTTTTaat taggaaaacg atttctgtga	10080
tcacttgtga aactggaggt tgatttgggg cataggacaa tagatttgat gctttgcaaa	10140
aagctgtttc aaagcagaga aatggaatag agacaattat gtagcgagga gggagggtgg	10200
ggcgaagatg gagacagaga agtggaagct gactttaggg aagaggaaca tagaccacag	10260
gggcggggcg gggggcaggg gcggggggcg gggctcaaag gaggcagtgg gaacgttgct	10320
agtgttcga gcgtaagcgt gaatgtgcaa gcgtctttgt ggtgtgtgac caggagtagc	10380
gtggctggct tgtgtgtgc ttgtaatecc agtctttgag gtttccacac gtttccacag	10440
tgggtgtgat ttccctcgg agagcatgag ggctctgctt tccccacatc ctccccagcg	10500
ttcgttgga tttgtttcca agatgttagt gggtagaca aagcctctct gttgatttgc	10560
ctttaacagg tgacaaaaa agctcaacca ggagacattt ttgccttctt ggaaggtaat	10620
gtcccatgt agagcaatgg gaccatctc taaggtgagg ctactcttgc agtttgcacc	10680
cagctcttct gatgcaggaa ggaagttagt gggcaagcaa gactgtttgc ttcttgcgat	10740
ggacacattc tgcacacaaa ggctcaggag gggagaaggc tgtttgatgt ttagcactca	10800

ggaaggcccc tgatgcatct gtgattagct gtctccatct gtggagcaga cacggactaa	10860
ctaaaaacca gtgttttttaa attgtcaagc ctttaagggtg aggaaattga cttattgtgc	10920
tgggccatac gtagagcaag tgctctgcat tgggccaaacc cccggctctg gtttctaggc	10980
accagaatgg cctagaacta actcacaatc ctcccatcc aggtctcagg tgctagaatg	11040
aaccactata ccagcctgcc tgccctgcta cctgccttcc taaattttaa atcatgggga	11100
gtaggggaga atacacttat cttagttagg gtttctattg ctgtgaagag acaccatgag	11160
catggcaact cttataaagg aaaacattta gttgggtggc agtttcagag gtttttagtac	11220
attgtcatca tggctgggaa catgatggca tgcagacaga catggtgctg gagaaaggga	11280
tgagagtcc t acatcttgca ggcaacagga cctcagctga gacactggct ggtaccctga	11340
gcataggaaa cctcacagcc caccctcaca gtgacatatt tccttcaaca aagccatacc	11400
tcctaatagt gccactccct atgagatgac agggccaatt acattcaaac tgctataaca	11460
ctttaaaagta ttttatTTTT attattgtaa attatgtatg tagctgggtg gtggcagccg	11520
aggtgcacgc ctttaatccc agcacttggg aggcagaggc agatggatct ctgtgagttc	11580
aagaccagcc tggctataaa gagctagtgt caaggaagga tatacaaaga acagttctag	11640
gatagccttc aaagccacag agaagtgtct tcttgaaaac caaaaattgt gctgggacct	11700
gtctctgctt tggttgcttc ccactcccc agagctggac tcttggtcaa cactgaatca	11760
gttgcaaaat aaactcctgg attcctctct tgtaacagga gcccgaaatc aggcgcccac	11820
ttgtcttctc gcaggattgc catagacttt ttctgtgtgc ccaccattcc agactgaagt	11880
agagatggca gtggcagaga ctgggaaggc tgcaacgaaa acaggaagtt attgcaccct	11940
gggaatagtc tggaaatgaa gcttcaaaac ttgcttcatg ttcagttgta cacagactca	12000
ctcccaggtt gactcacagc tgtaaatatt cctgactatg tctgactgc ttttatctga	12060
tgcttccttc ccaaaatgcc aagtgtacaa ggtgagggaa tcacccttgg attcagagcc	12120
cagggtcgtc ctcttaacc tggacttgct tttctccggc agcctctgac accctcccc	12180
ccattttctc tatcagaagg tctgagcaga gtgggggcac gctcatgtcc tgataactc	12240
cttgtcttcc tgaagatcta acttctgacc cagaaagatg gctaaggtgg tgaagtgttt	12300
gacatgaaga cttggtctta agaactggag caggggaaaa aagtcggaagc tggcagcatg	12360
taccgaaat ccagaaactg gggaggtaga gacggaatgag tgcccggggc tagctggctg	12420
ctcagccagc ctagctgaat tgccaaattc caactcctat tgaaaaacct ttaccaaaaca	12480
aacaacaaaa caaataataa caacaacaac aacaacaaac taccatatac aaggtgggag	12540
gtctttggct cttgaggaat gactcaccca aacccaaagc ttgccacagc tgttctctgg	12600
cctaaatggg gtgggggtgg ggcagagaca gagacagaga gagacatgac ttcctgggct	12660

gggctgtgtg ctctaggcca ccaggaactt tcctgtcttg ctctctgtct ggcacagcca 12720

gagcaccagc acccagcagg tgcacacacc tcctccgtg cttcttgagc aaacacaggt 12780

gccttggtct gtctattgaa cgggagtaag ttcttgacaga tgtatgcatg gaaacaacat 12840

tgtcctgggtt ttattttctac tgttgtgata aaaaccgggg aactccagga agcagctgag 12900

gcagaggcaa atgcaaggaa tctgcctcc tagcttgctc cccatggctt gccgggcctg 12960

ctttctgcaa gcccttctct ccccatggc atgctgaca tgaacagcgt ttgaaatgct 13020

ctcaaatgtc actttcaaag aaggcttctc tgatcttgct aactaaatca gaccatgttt 13080

caccgtgcat tatctttctg ctgtctgtct gtctgtctgt ctgtctatct gtctatcatc 13140

tatcaatcat ctatctatct atcttctatt tatctaccta tcattcaatc atctatcttc 13200

taactagtta tcatttattt atttgtttac ttactttttt tatttgagac agtatttctc 13260

tgagtgcagc ccttggtgtg cctggaacct attctgtaac cagctgtcc tcaaactcac 13320

agagatccaa ctgcctctgc ctctctgggtg ctgggggttaa agacgtgcac caccaacgcc 13380

ccgtctatc atctatttat gtacttatta ttcagtcatt atctatctc taactatcca 13440

tcctctgtct atccatcatc tatctatcta tctatctatc tatctatcta tctatcatcc 13500

atctataatc aattg 13515

<210> 2

<211> 14931

<212> DNA

<213> Cricetulus griseus

<220><221> misc_feature

<222> (2716)..(2239)

<223> n is a, c, g, t or nucleotide is missing

<400> 2

catgtacact tatgcaagta tgatatggcc caacacagta ttttacacca atttttatct 60

ataaaatata catgtacatc aaaatatatt attaataata acatcattat tctttctttc 120

caagtaataa acacatacac tgaaattttg gttcttgtgg ataattttaa tgaaacagga 180

aatgcaaatt tatcttagca tgtttacttc actttctttg catagataac cagtaatcac 240

attgatggat catgtagtga aatgtatttt taggtatcta aggaattttg gcttcgtttt 300

gigcttgttg aactgaatt ctattcctaa caacagtgtg taaggattct gtctgatttc 360

ttttaccagt attgtccat ttgcattttc ttattatttc atggetgctg ttctagaaag 420

tggaaggtag tgtgtcaagt ctgtttaaca tgtttccctg atgatcagtg tcttaacacc 480

tctctgagta catgttggcc aatgtcgttt ctgacccat ctattcttgc ttgacttate 540
ctggtacatg cctgccaaga aatttctcct catcctttct gtctcttcac tgatttactt 600
gatgtgtgga ttccacattg atcatatgga aatagaagat acaattttct ttattcacag 660
tttgggaagac ttccaatctc atagatcatc attatTTTT gctactgttc cctatgctat 720

ggtgaaattt ccatttgaat aattgcttaa acaattaaca agaaagaatc tatttttact 780
tgcaataact tccatttcag aacatttact aactgttac tataatccaa aactagtttt 840
atatacatg tgagaaatga ctaattcata atttggccat gacatttttt tcagaaacag 900
aaaaagtac caatacatc acaatgctat aaatattaag acttcagcaa attaaatatt 960
tattcatgat atcacataaa attcatttat tatgttttat ttaaatgtgt ttttaaaca 1020
gtggtatcac taaatattaa gtttagatgtg tttatgtgct taatgaattt atattttaga 1080
atgttataag ttgtatatag tcaaatatgt aataaatttt attttttagg tctttctcat 1140

taaggatatt taattttggg tcccttttcc agagtgactc tagctcatga tgagttgaca 1200
taaaaactaa acagtacaaa atgtacattg cattcagtat tgcacttgat ctttgcactg 1260
aagtttgagt cagttcatac atttagtact tgggaagtac attaagctaa ctttcattgc 1320
tctggcaaaa tgctcgataa gataagagtc tatttgggaa agccatggca gcaggaaagt 1380
aagactgctg atgatgttta atccatagtc aagacgcaga aggagatgaa tgctggtatc 1440
caacattttt tgctgttcat tttctctaga accctagtcc ataaagatgt atgacttgca 1500
ttcaaaatgc gtcccttca gttgttcaac ttttctgtaa atatccttc aggcatgtct 1560

agaagattgt ttgcaaaata cttctcaatc cattcaagtt gatagtgcag attaatcact 1620
gcagaataaa agcctgtaac ttggctcacg tgccaaggaa tatgcacact cctgacacat 1680
caataagtaa atcaaagtg agcttttggc ttaacattg ccagacttat gtaatgttct 1740
gcacgttctt cctccatcac tttttattct aatggtgttt cttgacatt gaatcacgt 1800
gtggaagctg cttagaatta acattgaaat ctactgatat atttatgatg cagcaattta 1860
gatttactat ttacttaga attttttata attgagagaa tataatattt tcacagtat 1920
ctatctgctg taaatagagg attttaaaaa aaatctctat aactttttt tacaacacac 1980

agtaaaatta agttaaaatt taataaagtc actatgttga tttcaaagtg tgctacgcc 2040
acggtgtgta cgcaggtgta gcagaagatg ccactaaggt gggctaaggc cgtgggttg 2100
gggtctgcgc tccctggaga tgagccccag gcggttccct ggcaatcagc tgcatcatg 2160
atgccgatg agccannnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2220
nnnnnnnnn nnnnnnnnc tgggtgactt tatggaaaga atttgataga ttcatgatg 2280
tagaagaatt ttattaggct tattttacag gagactaaga cctgggacc taaagatgc 2340

tggtcctga gaatcaggaa atgggtagag acgtggttga tggatgaga cagattttag	2400
agaactctta gatcatgggc aatgaccgca atctgatgct tagaatagat catctataaa	2460
caattatgct gttctttttc tttctgttgt atgatctgat gatgtagccc ccttgccaag	2520
ttccctgac ccccttgcca agttccctga ttgtaacagt atataagcat tgcttgagag	2580
catattcaac tacattgagt gtgtctgtct gtcatttctc cgccgattcc tgatttctcc	2640
ttgagccttt tcccttgctc tccctcggtc ggtggtctcc acgagaggcg gtccgtggca	2700
aaagtgtata aatgttctaa aacatttgaa ctctaaaaca tgcaaatga aaaattaaaa	2760
taaataaaca tgaaaattaa aatatattag ctgctaaaag ttaaacaata ctatataata	2820
ttttgttatt agaattcaaa atcacattag ttggatttaa ttgaaacatt gcattctttc	2880
aataataatt tcaataaaaa aagtttcccc atgatagtag aaaataataa catatgtatc	2940
tatctattta tttaactaca catatatagc atttgtttca actaaaataa atgaatgagc	3000
aaagcaccta agtaattggt gtctattata tttatgaagc caatagtttc aaataaatta	3060
tcatgcataa ggaggtattg caaatgttaa accttttttg aaacagatat tcccagttac	3120
agaaattata atttctaate tttctataa gtagaatgat gataattaat ataggccatt	3180
tgtaaataat gttcagatta aaatattctc tatttcacta gagaagaatg atattaaatg	3240
tattatattt tatttcccat ttgttttga ccactattct atatccctca gcagtttaaa	3300
tttgtttcac catatgtgtg tgtgtttgia tcttaaatat ggcaactaaa ttagaataat	3360
ttaataataa tctttaggag aaaagatatt gaattatttt atgttgatag gaaaatatct	3420
tttaattgtc caagaatact ttttcttcta ttttaggact gatcagaccc aggactaata	3480
ttttatatgt actaattcta tgtacaaaa tatgttatta tctcatgaat tctgtctcaa	3540
tattgaggta ataaaaatg tccatcatga actttaaaat taaaataatg attaattaat	3600
ttttattcat atttgtttg tatgaatggt tatacatcac atgtgtgcct ggtgactgtg	3660
aatgtcagga gaaggatga aagccactgg aattggaata agagataata ttgagatgt	3720
tatgtgggtg ctgagaatta gacgcaagcc atcttcaaga atagccagca tactatacca	3780
ctgagtaatc cattcatccc tcaataatta tctttgtaga cagtaaatat atttctaaac	3840
tataaatgac cagaaaaatt aatgtattat taatgaagac attcatctca tgtgacacac	3900
ttcacctgtc taaatcagta acactctctc cactaattaa gattttctaa gtgcatgaca	3960
cttactatth cttaaagctgt ccaatggggg ccagtcccca gtcagcacc agtgagataa	4020
tccatgaatg catttatatc ttaggaaaaa ttcttatcta tgtagtattt agaacatttt	4080

catgtgaggg gataaacaag gaagcacaga tgctttctga tagaaacttt ctctttaatt	4140
catctagaaa aaaaaaacct ctcaggaaaa tctctcttgc tctcctccca atgctctatt	4200
cagcatcttc tccctactta attctagatc tttttctcta tgcctccttg ctgctgccct	4260
gtggctctg ctctatgcct ccccatgtca cttttctttg ctatctcacc gttaccttct	4320
ctgcctcact ctctgccttc ttctctgctt ctcacatggc caggctctgg acaattatag	4380
ttatatgtta cattctcata acacatgata tgtcacatag tttctctcag gctagggata	4440
tcacaatgac tggccaatga gcaagtggcc ttgcatgtag ctctaagttg gtgatggttc	4500
ccagacagta agtagccatt tggttgaaat ttgaggttgg gtagtacatg aagactgaat	4560
tttcttcaaa ctctggcctt gaaatagtaa aacaacacct atgaaaatga cgacctgtat	4620
ttgtcttttag aggcaaccac atattgtctg cagggcctgc tttgaatttg ctctgaagtt	4680
agcttgtttg tgtaaaagga agaatcctat atcagcctga gaaatgtaaa atatcctagc	4740
atttcaagtc atcaaaatta tatggagagt ataaatcatc cttctgacta ttcatagtca	4800
tatttgtgtc caccaagtat aaaacacact accaaagggc tgtggaaaaa atcgccataa	4860
ctgttcttat tagggaggca tagcagtggg acctgaggaa gttacagcaa caaccagtca	4920
tccagtcaat aaccccatgg ctttgccact tggaggtacc caataatgtt tggctttgcc	4980
gagtaggact ccaacaaatt cagaggggtca atttttaaat gctggttgtc actgctgaac	5040
agtccattg cctctgcat aattccacaa tggaaagctt tttacactga ttgccaatca	5100
ttaaacagcc tactcagcat aaacaggtat gatattattc tgcattttgt tacattacta	5160
gatgaattcc tatttcttcc tacaatagtg gaactgaaaa aagatacaca atcatactac	5220
ccctctacta atcttatgac ttatatcatt tcaattttca gaccataatg caaactattg	5280
acaaaacat gtgaagatga aaaatagaaa tgtagaataa tattacatat aaaaagaaaa	5340
ggcggactta ttttgtttta tttcttagca tgcatagcaa tacatgattt gaggtttata	5400
taataaaggg acaataaatc ttcaagaaac ttaccctac tgaattaaaa tattaagaa	5460
ggtcacacat ttactcaaat atattagact actgggcaaa tagacatgaa aagtagagtt	5520
aatattgagg taggccttct gtgaaatgtc taaggaaatt atgtttcata cagtgtgtaa	5580
ccaagtggga atcatatcag aaagcagtca aaagcttata ttacaagtaa cagatgcttg	5640
gttatatgac ctccagagc ttgactgtct atacacaaaa agtggtgtta ataaaactgt	5700
aatttgggct atgttttttt aaatggcttc accaacaatga aaggaaggga atgagcatgt	5760
catggatgct tagagattat gcttcagca agaagaattg agctttggct cttattacag	5820
aaacatgaca aggtgtgagt tttatttatt agaaattata taatatitaa agctggggac	5880
taaaaatit atitgaaaca acaggcaagg gataggcatg tactagaagc aaaaatagga	5940

tgtcaatgct gtaatgttat tttttggacc aaaatagtat ttcctataga aatgacaatg	6000
atcttaggtt attattcttc ataaagatga caagttcaca agatataccta gttcattaaa	6060
atcgtttttag tcatTTtaata gagtgcTgtg atagattaca caaaggaaag cacttacgat	6120
gagaaataat gatataccaca attatTTTct taattcttag aaacattcta ttgttatatc	6180
tcaatctcag aagccactta ttgctttatt attgaaacat atgaaattgt aagttatata	6240
ttgtctatgg tgacatttca aagaacatgt gacgtacagt gtagcacaga taaagaacat	6300
aactgcagct gaatcagtaa ctaaacttac atacattaaa tctgccatgt tggcaacagt	6360
gtgtgcacta ccaaaggatg tactaatgct cagcacactc ccctatgtca ccctttgttc	6420
atcattacat cataggctca ttttgtttgc ttttgaaatc tagaccaagt cttttgtgtc	6480
ttccaagca cagagctcat taattacct catagacttg ttaaacttct tctggttcat	6540
caattgaata gaaatactca ctactaatta tgtgagaccc tgccagtacc atagcacatg	6600
gataatTTTT acataaaaca tgcatacaag taagattatt cagactgaac atgaatTTta	6660
gagaaatcag gaaggagtat atgggagtgg ttggagtgag actagagaaa tgtaattaaa	6720
ctataatctc aatacaaaaga tctactaagc aaaaaacatg aaacattgtc attcaagtga	6780
aacatcagtc ttcaaatTgg aaagatatTT ttactaggaa aatgtctggt agatggttat	6840
tatctagaaa acacaaaaat tagaaaacgg taaactTTaa taaaaagaat aatacaatga	6900
gactacatga aaagtTctta actaatgaaa caaatatctt gaaactTTTT tcttaaaagt	6960
ttaatataca taaccatcat ggaaattcaa attaaaacta tttacatat acccctgaaa	7020
taataactaa tacccaataa aaataatata aacaaaaaat ggcaatgcat gccatcatgg	7080
atttgggaga gagaatgttc attgcagttc tgaatggata ctggtgccac cacggtgaaa	7140
atctctgtat aggtccttcc aaaagctgaa aatagacata tcacaagacc tgccacacat	7200
ttttcaagca aatacccaaa ggactctacc tgactgcaga gacactttct cataaaatat	7260
tattgttgat ctattcataa tatctggaaa atagaaacag ccaagatgcc catcaactga	7320
ttaatagatg ataaaattat tgtacatttc agtgtaatat tattcagttt ttaagaaaaa	7380
tgaaattatg taataagcat gtaaatggat atatcttgaa acaaccattc cccattatat	7440
tacctaaca ttgaaagtc aaaatcata gatctTTTT gtggatctac taatcttttg	7500
ctatatgtat tttattgaac taccatgga tgtgagataa ttggtacaa cagcacatgg	7560
gagagcatgg gatcattcaa ggaagattag agagaatgca ttttttagga gataatggag	7620
gagcaataga aaggattaaa tgaggttact gatgaaagt atggttagag aaggcaatat	7680
gaggagggat aactagcact tagggccttt tgaaaaagac atagagaaaa tactattgta	7740
gaaacttctt ataattggtg tatagttata tacaccaaag agctcagatg gagttaccct	7800

ataatggaaa tattaactac tttttatcac tgtgataaaa catcctgaac agagcaacat	7860
agattgggaa gcatttactt tggcttacag ttctaacggg ataaaaattc atgatgaaag	7920
aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga agcaggtgag agctcacatc	7980
ttgaagtgtg agaattgtagc agagagaaca aactgcaaat gaccagaaaa tgcttttggg	8040
tcagagccca taccctctg actgacttct ccagaaattc tgaacaaata aaactcccca	8100
aacagagcca taactgaagg tccagtgtct gagactacta ggggtatttc ttattcaaac	8160
cactacaatg gggtaggggg agcaatctc caagtaggca ctacacacag acaataaaaa	8220
actctagtaa ctggaatgga ttgacttatt tgaattactt gccagtggag ctacatagag	8280
cacaattatt gtattttaa taccctttat gatcttaca aacttgacag taagatcata	8340
ttgctaaaga aaccacatat ttgaatcagg gaacatgggtg atatctagtt gttcttcaac	8400
tggaaacttc atgctttctg ccagcattc atgttgctgg aaagagcaat gtacactacc	8460
agtgtagaaa ttaaatcatc aatcttatca agatgtggat cctataagtt acaataaaaa	8520
ttagcctgat aagatatccc caccagaaga atattcacat aaatgctatg ggagcaacaa	8580
gctattttct aaattagctt taatcctatt ctacaagaga gaatccatat ctagaatagt	8640
tatagggatc aagaacccat ggcttgattg gtcataggcc caatgggaga tcctaattatt	8700
attgttctac aaaatgaaaa taactcctaa tgacttggtg ctgcagtaat aagttagtat	8760
gttgctcaac ttcacaaga gaagttttgt cttacaataa atggcaatta aagcagcccc	8820
acaagattta tatcataccg atctctcat ggcctatgca tctagaagct aggaaacaaa	8880
gaggacccta agagagacat acatggtecc cctggagaag ggggaagggg caagacctcc	8940
aaagctaatt gggagcatgg gggaggggag agggagttag aagaaagaga aggggataaa	9000
aggagggaga ggaggacaag agagagaagg aagatctagt caagagaaga tagaggagag	9060
caagaaaaga gataccatag tagagggagc cttgtatgtt taaatagaaa actggcacta	9120
gggaattgtc caaagatcca caaggtcca ctaataatct aagcaatagt cgagaggcta	9180
ccttaaaagc ctttctctga taatgagatt gatgactacc ttatatacca tcctagagcc	9240
ttcatccagt agctgatgga agcagaagca gacatctaca gctaaacact gagctagtgtg	9300
cagacaggga ggagtgaiga gcaaagtcaa gaccaggctg gagaacaca cagaaacagc	9360
agacctgaaa aaaaattgac acatggaccc cagactgata gctgggagtc cagcatagga	9420
cttttctaga aacctgaat gaggatatca gtttgagggt ctggttaatc tatggggaca	9480
ctggtagtgg atcaatatct atccctagtt catgactgga atttgggtac ccattccaca	9540

tggaggaatt ctctgtcagc ctagacacat gggggagggt ctaggtcctg ctccaaataa	9600
tgtgttagac ttgaagaac tcccttgaga agactcacc tccctgggga gcagaaaggg	9660
gatgggatga gggtttgtga gggacaggag aggaggggag ggtgagggaa ctgggattga	9720
caagtaaattg atgcttgttt ctaattttaa tgaataaagg aaaagtaaaa gaagaaaaga	9780
aaacaggcca aaagattata aaagacagag gtggtgggtg actataaaga aacactatta	9840
tctaaataaa aatatgtcag aagcacacat gaacttatag tgtttatgaa agtatgtata	9900
ataactacat aatctcaagc caagaaaaaa atatcatctt tcagtgtatga aggtgatattt	9960
atttctccca gaattaaagc caaagaccta atgaaagtaa ttatcttcaa aaggttgaaa	10020
atacatactt tgcaatacac agatctgcct agaaatctca tgttcacaat acacatgatg	10080
ctcaattgaa ttccattcaa tgttacagtt tagataaaca gttttagat aaactcacia	10140
tgtatcatctt ctttttattt ttgacccaaa cagcttctca tctgttatc agaataattc	10200
ctcgatggca ggatatccat cccaattggg ggaaggggag aatttgaaga aaacctagac	10260
cacatacata ttgccattg ggaaacaaag tctaaaatga tgttgttcac atcttctcta	10320
ctagtcctct ccccgctcca aagaaccttg gtatatgtgc ctcatcttac agagagagga	10380
aagcaggaac tgagcatccc ttacttgcca tctcaaccc aaaatttgca tcattgctca	10440
gctctgccct tctcatatga cagttacaag tcaaggcttc caaagtcct ctgtcatgtt	10500
tgggtgcaat agtttataca gatgacttca tgtcttcata tctaatgtct tatatagatt	10560
aatattaaac aatgttattt ctctaaccac attttaatt aatttaaaaa tccattaatt	10620
gtgtctataa aatgcagaca gagtgtgag acacaatata agcctgatga tctgaatttg	10680
aaactcacac ccaccacatg gagaatcaac ttccaaaaat tttctatta ctccacact	10740
tacaccattg tacaaacaca ataataatga acaaaatgaa atgaaataaa aaattaagtc	10800
tctgtaggta atgctactgt gcagcaaaag taaaaatggc agcttaagct tgctttatgg	10860
ttacacttta ccatcttcca ttaattataa ggacttcaat catggcagaa ctatgtgtt	10920
attgtctcag tgtaacctaa ccagggtgtc cagatgttct taatgtggac acctaaacta	10980
tttgatattt gggtaagat ctttcctct ttcagaagaa acctcaggac agagggaatc	11040
ttgtcttta attttgagtc tgtagacttt ttccatttca aatatacatg aaacaagtga	11100
tgaagaaat taatcaaaag gtgggaattg caatgatatt aggttcaata ttaagcttca	11160
atattatcat ggaatcgct gttatacact gagtgtttgg caataaggga tttttagaag	11220
aaggagtttt tattctcaac aggttcctta agtttagctc aaataaatct aagcaatcca	11280
ctctagaatt aaatagtttc ctaagggcac agctatgaat agagctcaat ttacatataa	11340
aattttgttc accatttatg tcatccagc ttcttagt acaaggaaaa tacaaaatat	11400

ttagatgtca atatcaagtg aatagttcat ctcccttttt aatatatc acctaaatca	11460
ccattttctc agaaaaatct ggcctgaagi tctgtctgga acttcaacat gaaaaatag	11520
cacagcttgc tattataaat cctagttgat ttttaagatt catgtctggt gtctgactca	11580
gaggggccag aggctagaca aatatTTTTT gaatcttcat tgtgaagatt tttaatgatt	11640
attttaatat aaataacaaa gatgatggat aatgtaactt tgtacagttc atagacgctg	11700
aactactttg tgcttaaaat gttagtcc taccataaat gataggtgat aagtgtatgt	11760
ttaatacttt ccctctgagc tataattcatg tactagagaa ttattttaaa catgaaaaga	11820
ctgtgtttat agtctcagct cctgagaact ggtccaacct taggcaggtg aatgccagga	11880
gcaacgtttt tcttctacag aggatgcttt gctgccaagc aacctggttg tgtggaaatg	11940
ttcctttttt aatcaagttt aaagggtctt catcatgctg ttgctccaca tattttcagg	12000
ttagagcttg gtccctggag tattatcttt taccagaaaa ttcatagtat tctttcaata	12060
actaacaact aaacttttcg ataaaaaga attggaattt caattttaaa gcctgagtaa	12120
aattcttgtg aatcaggata ttttatttta agtcttatct tttaaaaagt tattttattt	12180
tttaaaaaat tataataaac tttcataatt tcctctcttc acttttcttt acaaacactt	12240
ctatagatca ccatgtgttt ttttttttac atttatggcc tctttctgtt cattgttatt	12300
acatacaaat agtcttgctt atagaagaac accacaattt gttacctgat aacaaattat	12360
caaccttaa aacctacaaa ctattgatat tactgaaaag actatactta tagatgtaaa	12420
gatatatgtg tgtgcacata tatagatata catatatgta ggatttttaa ttttagattt	12480
tagacatcaa aattatttat atgactgaga aactagacac tataaatgag cattcagtat	12540
tcaacaccgt gattttagat attgtcacia tgacagaaaa ttttcttata gaaaatttta	12600
agttttgtga ttgctctgtg cacttagtga agtctcacag aaaaagaatc atagtatttt	12660
tagtttataa taaaaagtac atataattaa aatggttggc acaaaacaac atttgagcat	12720
ttttcctatt tactatcaag tagtatcatt ttgaaataat aatttgacta gtttcaaaaa	12780
tgaaaacaaa atttaaaact aatgcctaat ctagcctgat aacattttta tgaatgaaat	12840
tattcaatag tgttatcaat taggggceca aaacttttcc taaaataaaa cttttaattt	12900
ttttccattt ttatttaaat tagaaacaaa attgttttac atgtaaatca gagtttcctc	12960
acctccctt tctccctgtc cctcaactac acctacttg tcccatacca tttctgctcc	13020
ccaggaggagg tgaggccttc catggggaaa cttcagagtc tgtctatcct ttcggatagg	13080
gcctaggccc tcacccattt gtctaggcta aggtcacaaa agtttactcc tatgctagtg	13140
ataagtactg atctactaca agagacacca tagatttctt aggttctctc actgacaccc	13200
atgttcatgg ggtctggaac aatcatatgc tagtttctta ggtatcagtc tggggacccat	13260

gagctccccc ttgttcaggt caactgtttc tgtgggtttc accaccctgg tcttgactgc 13320

tttgctcacc actcctccct ttctgtaact gggttccagt acaattccgt gtttagctgt 13380

gggtgtctac ttctactttc atcagcttct gggatggagc ctctaggata gcatacaatt 13440

agtcacatc tcattatcag ggaagggcat ttaaagtagc ctctccattg ttgcttggat 13500

tgtagttgg tgcacatctt gtagatctct ggacatttcc ctagtccag atatctcttt 13560

aaacctacaa gactacctct attatggat ctcttttctt gctctctct attcttccag 13620

acaaaatctt ccgtctccct tataatttcc tctccctcc tcttctccc ttctcattct 13680

cctagatcca tcttccctc ccccatgctc ccaagagaga tgttgctcag gagatcttgt 13740

tccttaacce ttttcttggg gatctgtctc tcttagggtt gtccttgttt cctagcttct 13800

ctggaagtgt ggattgtaag ctggtaatca ttgtctccat gtctaaaatc catatatgag 13860

tgatgtttgt ctttttcta ctgggttacc tcaactcaaaa tggtttcttc catatgtctg 13920

tggatttcaa tagcacaaac aacatacagt atcttggggc aacactaacc aaacaagtga 13980

aagaccagta tagcaagaac tttagattta aagaaagaaa ttaaagaaga taccagaaaa 14040

tggaaagatc tcccatgctc tttagataggc agaatacaaca tagtaaaaat ggcaatcttg 14100

ccaaaatcca tctacagact caatgcaatc cccattaat accagcacac ttcttcacag 14160

acctgaaaga ataacttta actttatatg gaaaaacaaa agaccagga taggccaac 14220

aacctgtac aatgaaggca ctccagagg catcccatc cctgacttca agctctatta 14280

tagagtaata atctgaaaa cagcttggta atggcacaaa aatagacagg tagaccaatg 14340

gaattgagtt gaaaacctg atattaacct acatatctat gaacacctga ctttgacaaa 14400

gaagctaagg ttatacaatg taagaagaa agcatcttca acaaatcgtg ctggcataac 14460

tggatgctgg catgtagaag actgcagata gatccatgct taatgccaatg cacaaaactt 14520

aagtccaat ggatcaaaaa cctcaacata aatccagcca cactgaacct catagaagag 14580

aaagtgggaa gtatccttga ataaattggt acaggagacc acatcttgaa cttaacacca 14640

gtagcacaga caatcagatc aataatcaat aatgggacc tctgaaact gagaagcttc 14700

tgtaaggcaa tggataagtc aacaggacaa aatggcagcc cacggaatgg gaaaagatat 14760

tcaccaatcc tatatctgac agagggtctc tctctatttg caaagaacac aataagctag 14820

tttttaaac accaattaat ccgattataa agttgggtag agaactaat aaagaattgt 14880

taacagagca atctaacttg gcagaaagac acataagaaa gtgctcacca t 14931

<210> 3

<211> 4001

<212> DNA

<213> Cricetulus griseus

<400> 3

ccaagatgcc catcaactga ttaatagatg ataaaattat tgtacatttc agtghtaatat	60
tattcagttt ttaagaaaaa tgaaattatg taataagcat gtaaatggat atatcttgaa	120
acaaccattc cccattatat tacctaaca ttgaaagtcc aaaatcatat gatcttttta	180
gtggatctac taatcttttg ctatatgtat ttatttgaac taccatgga tgtgagataa	240
ttggttaaaa cagcacatgg gagagcatgg gatcattcaa ggaagattag agagaatgca	300
ttttttagga gataatggag gagcaataga aaggattaaa tgaggttact gatgaaagtg	360
atggttagag aaggcaatat gaggagggat aactagcact tagggccttt tgaaaaagac	420
atagagaaaa tactattgta gaaacttcct ataattgggtg tatagttata tacaccaaag	480
agctcagatg gagttaacct ataattggaaa tattaactac tttttatcac tgtgataaaa	540
catcctgaac agagcaacat agattgggaa gcattttactt tggcttacag ttctaacggg	600
ataaaaattc atgatgaaag aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga	660
agcaggtgag agctcacatc ttgaagtgtg agaattgtgc agagagaaca aactgcaaat	720
gaccagaaaa tgcttttggg tcagagccca taccctctg actgacttct ccagaaattc	780
tgaacaaata aaactcccca aacagagcca taactgaagg tccagtgtct gagactacta	840
ggggtatttc ttattcaaac cactacaatg ggggtggggg agcaatcctc caagtaggca	900
ctacacacag acaataaaaa actctagtaa ctggaatgga ttgacttatt tgaattactt	960
gccagtggag ctacatagag cacaattatt gtattttaat taccctttat gatcttaca	1020
aacttgacag taagatcata ttgctaaaga aaccacatat ttgaatcagg gaacatggtg	1080
atatctagtt gttcttcaac tggaaacttc atgctttctg ccagcattc atgttgctgg	1140
aaagagcaat gtacactacc agtgtagaaa ttaaatcacc aatcttatca agatgtggat	1200
cctataagtt acaataaaaa ttagcctgat aagatatccc caccagaaga atattcacat	1260
aaatgctatg ggagcaacaa gctattttct aaattagctt taatcctatt ctacaagaga	1320
gaatccatat ctagaatagt tatagggatc aagaacccat ggcttgattg gtcataggcc	1380
caatgggaga tcctaattatt attgttctac aaaatgaaaa taactcctaa tgacttgttg	1440
ctgcagtaat aagtttagat gttgtctaac tctcacaaga gaagttttgt cttacaataa	1500
atggcaatta aagcagcccc acaagattta tatcataccg atctcctcat ggcctatgca	1560
tctagaagct aggaacaaaa gaggacccta agagagacat acatgggtccc cctggagaag	1620

gggaaggggg caagacctcc aaagctaatt ggagcatgg gggaggggag agggagttag 1680

aagaaagaga aggggataaa aggagggaga ggaggacaag agagagaagg aagatctagt 1740

caagagaaga tagaggagag caagaaaaga gataccatag tagagggagc cttgtatgtt 1800

taaatagaaa aciggcacta gggaattgtc caaagatcca caaggtccaa ctaataatct 1860

aagcaatagt cgagaggcta ctttaaaagc ctttctctga taatgagatt gatgactacc 1920

ttatatacca tcttagagcc ttcattccagt agctgatgga agcagaagca gacatctaca 1980

gctaaacact gagctagttg cagacaggga ggagtgatga gcaaagtcaa gaccaggctg 2040

gagaaacaca cagaaacagc agacctgaaa aaaatgttgc acatggaccc cagactgata 2100

gctgggagtc cagcatagga cttttctaga aacctgaat gaggatatca gtttgagggt 2160

ctggttaatc tatggggaca ctggtagtgg atcaatattt atccctagtt catgactgga 2220

atttgggtac ccattccaca tggaggaatt ctctgtcagc ctagacacat gggggagggtt 2280

ctaggtcctg ctccaaataa tgtgttagac ttgaagaac tcccttgaga agactcaccc 2340

tccctgggga gcagaaaggg gatgggatga gggttggtga gggacaggag aggaggggag 2400

ggtgagggaa ctgggattga caagtaaatg atgcttgttt ctaattttaa tgaataaagg 2460

aaaagtaaaa gaagaaaaga aaacaggcca aaagattata aaagacagag gtggtgggtg 2520

actataaaga aacactatta tctaaataaa aacatgtcag aagcacacat gaacttatag 2580

tgtttatgaa agiatgtata ataactacat aatctcaagc caagaaaaaa atatcatctt 2640

tcagtgatga aggtgatattt atttctcca gaattaaagc caaagaccta atgaaagtaa 2700

ttatcttcaa aaggttgaaa atacatactt tgcaatacac agatctgcct agaaatctca 2760

tgttcacaat acacatgatg ctcaattgaa ttccattcaa tgttacagtt tagataaaca 2820

gttttgatag aaactcaca tgtatcatct ctttttattt ttgaccaa cagcttctca 2880

tctgttattc agaataattc ctcatggca ggatatccat cccaattggg ggaaggggag 2940

aatttgaaga aaacctagac cacatacata ttgccattg ggaacaaaag tctaaatga 3000

tgttggtcac atcttctcta ctagtctct ccccgctcca aagaaccttg gtatatgtgc 3060

ctcatcttac agagagagga aagcaggaac tgagcatccc ttacttgcca tctcaaccc 3120

aaaatttga tcatgtctca gctctgcct tctcatatga cagttacaag tcaaggcttc 3180

caaagtcct ctgcatgtt tgggtgcaat agtttataca gatgacttca tgtcttcata 3240

tctaattgtc tatatagatt aatattaaac aatgttattt ctctaaccac attttaaatt 3300

aatttaaaaa tccattaatt gtgtctataa aatgcagaca gagtgtgag acacaatata 3360

agcctgatga tctgaatttg aaactcacac ccaccacatg gagaatcaac ttccaaaaat	3420
tttcctatta ctccacact tacaccattg tacaacaca ataataatga acaaaatgaa	3480
atgaaataaa aaattaagtc tctgtaggta atgctactgt gcagcaaaag taaaaatggc	3540
agcttaagct tgctttatgg ttacacttta ccatcttcca ttaattataa ggacttcaat	3600
catggcagaa ctatgctgtt attgtctcag tgtaacctaa ccaggtgttc cagatgttct	3660
taatgtggac acctaaacta ttgatattt gggttaagat ctttcctct ttcagaagaa	3720
acctcaggac agagggaatc ttgtctttta attttgagtc ttagacttt tccatttca	3780
aatatacatg aaacaagtga tgaagaaaat taatcaaaag gtgggaattg caatgatatt	3840
aggttcaata ttaagcttca atattatcat ggaatcgctt gttatacact gagtgtttgg	3900
caataaggga tttttagaag aaggagtttt tattctcaac aggttcctta agtttagctc	3960
aaataaatct aagcaatcca ctctagaatt aaatagtttc c	4001